

Université de Montréal

**Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes
lexicaux et analyse des stratégies de résolution de problèmes lexicaux**

**par
Dominic Anctil**

**Département de didactique
Faculté des sciences de l'éducation**

**Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de M.A.
en didactique**

Août 2005

© Dominic Anctil, 2005



LB

5

U57

2005

V. 033



AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

**Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes
lexicaux et analyse des stratégies de résolution de problèmes lexicaux**

présenté par :

Dominic Ancil

a été évalué par un jury composé des personnes suivantes :

Isabelle Montésinos-Gelet
présidente

Pascale Lefrançois
directrice de recherche

Alain Polguère
codirecteur

Marie-Claude Boivin
membre du jury

RÉSUMÉ

Cette recherche vise à déterminer 1) quels sont les problèmes lexicaux des étudiants universitaires en langue première et 2) quelles sont les stratégies de résolution déployées par les étudiants confrontés à des problèmes lexicaux.

En nous appuyant sur des concepts linguistiques de la lexicologie Sens-Texte, nous avons développé une typologie structurale des problèmes lexicaux à l'aide de laquelle nous avons classé les problèmes lexicaux relevés dans un corpus de textes d'étudiants. Cette analyse a révélé que près du tiers des problèmes sont attribuables à une méconnaissance des propriétés de combinatoire des unités lexicales.

Huit entretiens métagraphiques au cours desquels les étudiants avaient pour tâche de repérer et de corriger des problèmes lexicaux nous ont permis d'analyser leurs stratégies de résolution. Nous avons constaté qu'environ 60 % seulement des problèmes ont été repérés. Par contre, lorsqu'ils ont été traités, plus de 80 % des problèmes ont été bien résolus et les stratégies observées étaient en général appropriées. Il ressort de notre étude que les étudiants gagneraient à connaître davantage certains phénomènes lexicaux et à utiliser le dictionnaire plus efficacement.

Mots-clés : *erreur lexicale, maladresse lexicale, typologie des problèmes lexicaux, théorie Sens-Texte, étudiants universitaires, entretien métagraphique, stratégies de résolution de problèmes, propriétés de combinatoire, lexique, didactique du français.*

ABSTRACT

This research aims to determine 1) what are the lexical problems of first language undergraduate students and 2) what are the strategies they use when solving such problems.

By using linguistic concepts of the Meaning-Text theory, we have developed a structural typology of lexical problems with which we classified the lexical problems discovered within a corpus of student essays. This analysis has revealed that nearly a third of the problems are attributable to a misunderstanding of the combinatorial properties of lexical units.

To analyse students' problem solving strategies, we conducted eight interviews in which they had the task of identifying and correcting lexical problems. We have observed that only about 60% of the problems were detected, although, once analysed, more than 80% were properly corrected and the strategies observed were generally appropriate. We can conclude from our study that students would benefit from both a better knowledge of certain lexical phenomenon and a more efficient use of dictionaries.

Keywords : *lexical error, typology of lexical problems, Meaning-Text Theory, undergraduate students, problem solving strategies, metagraphic comments, combinatorial properties, lexicon, French language teaching, vocabulary teaching.*

TABLE DES MATIÈRES

IDENTIFICATION DU JURY.....	ii
RÉSUMÉ	iii
ABSTRACT	iv
TABLE DES MATIÈRES.....	v
LISTE DES TABLEAUX	ix
LISTE DES FIGURES.....	x
LISTE DES CONVENTIONS D'ÉCRITURE	xi
REMERCIEMENTS.....	xiii

INTRODUCTION	1
--------------------	---

I. PROBLÉMATIQUE.....	2
1. CONTEXTE DE LA RECHERCHE	3
2. PROBLÈME GÉNÉRAL	4
3. ANALYSE DU PROBLÈME.....	5
3.1 LE LEXIQUE AU POSTSECONDAIRE	5
3.2 LA MAÎTRISE ET LA COMPÉTENCE LEXICALES	7
3.3 LE LEXIQUE AU-DELÀ DE L'ORTHOGRAPHE LEXICALE	7
4. PROBLÈME SPÉCIFIQUE	9
5. OBJECTIFS DE RECHERCHE	10
6. PERTINENCE ET VISÉES DE LA RECHERCHE	10

II. CADRE CONCEPTUEL.....	11
---------------------------	----

1. LE LEXIQUE	12
1.1 NOTRE CADRE THÉORIQUE : LA THÉORIE SENS-TEXTE	12
1.1.1 LA LANGUE SELON LA THÉORIE SENS-TEXTE	13
1.1.2 LA LEXICOLOGIE EXPLICATIVE ET COMBINATOIRE	14
1.2 DEUX MISES AU POINT TERMINOLOGIQUES	15
1.2.1 LA DICHOTOMIE <i>LANGUE</i> ~ <i>PAROLE</i>	15
1.2.2 LA DICHOTOMIE <i>LEXIQUE</i> ~ <i>VOCABULAIRE</i>	15
1.3 LE TERME <i>MOT</i> : SOURCE DE CONFUSION	16
1.3.1 LES TROIS DIMENSIONS DU SIGNE LINGUISTIQUE	16
1.3.2 LE LEXÈME COMME REGROUPEMENT DE MOTS-FORMES	17
1.3.3 ET <i>MONTER SUR SES GRANDS CHEVAUX</i> ?	18
1.3.4 DÉFINITION DE LA NOTION D'UNITÉ LEXICALE (<i>LEXIE</i>).....	19
1.3.5 LES LEXIES AYANT UN SIGNIFIANT IDENTIQUE	19
1.3.6 LA SYNONYMIE ET LA PARAPHRASE	20
1.4 LE SENS ET LA COMBINATOIRE SÉMANTIQUE	21
1.5 LES PROPRIÉTÉS DE COMBINATOIRE	22
1.5.1 LA COMBINATOIRE SYNTAXIQUE OU RÉGIME	23
1.5.2 LES AUTRES PROPRIÉTÉS DE LA COMBINATOIRE GRAMMATICALE	24
1.5.3 LA COOCCURRENCE LEXICALE RESTREINTE : LES COLLOCATIONS	25
CONCLUSION	26

2. CARACTÉRISATION DE LA NOTION D'ERREUR LEXICALE	27
2.1 L'ERREUR ET LA NORME	27
2.2 LA PRÉSENTE RECHERCHE ET LA NORME	30
2.3 LE LEXIQUE ET LA NORME	31
2.3.1 LES DIFFÉRENTS VISAGES DE L'ERREUR LEXICALE	31
2.3.2 LE LEXIQUE ET LA NORME PRESCRIPTIVE	33
2.3.3 LES DIMENSIONS MOINS « PRESCRIPTIBLES » DU LEXIQUE	37
2.3.4 UN AUTRE TYPE DE PROBLÈME LEXICAL : L'UTILISATION D'UNE LEXIE FICTIVE.....	41
CONCLUSION	44
3. LES STRATÉGIES DE RÉOLUTION DE PROBLÈMES LEXICAUX	45
3.1 CARACTÉRISATION DE LA NOTION DE PROBLÈME	45
3.2 LA RÉOLUTION DE PROBLÈMES.....	46
3.3 TRANSPOSITION DE CES CONCEPTS AU LEXIQUE.....	48
3.3.1 LE PROBLÈME LEXICAL	48
3.3.2 REMARQUES TERMINOLOGIQUES.....	48
3.3.3 LES STRATÉGIES DE RÉOLUTION DE PROBLÈMES LEXICAUX	49
CONCLUSION	52
4. LE MÉTALANGAGE.....	53
4.1 REMARQUES TERMINOLOGIQUES.....	53
4.2 LE MÉTALANGAGE ET L'ACTIVITÉ MÉTALINGUISTIQUE	54
4.3 L'ACQUISITION DU MÉTALANGAGE : UNE FINALITÉ DIDACTIQUE?	55
CONCLUSION	56
III. MÉTHODOLOGIE.....	57
1. RAPPEL DES OBJECTIFS	58
2. OBJECTIF 1 : DESCRIPTION DES TYPES DE PROBLÈMES LEXICAUX OBSERVÉS	58
2.1 CORPUS D'ANALYSE ET ÉCHANTILLON	58
2.2 LES TEXTES DU CORPUS	60
2.3 LA CORRECTION DES TEXTES	60
2.4 LES SOURCES DE RÉFÉRENCE	61
2.5 L'OUTIL D'ANALYSE : LA TYPOLOGIE DES PROBLÈMES LEXICAUX	63
2.5.1 TYPOLOGIE DESCRIPTIVE OU EXPLICATIVE?	63
2.5.2 PRÉSENTATION DE LA TYPOLOGIE ET LIMITES DE L'OUTIL	64
2.5.3 ERREURS ET MALADRESSES : UN DEUXIÈME AXE D'ANALYSE	67
2.6 CARACTÉRISATION DU PROBLÈME EN FONCTION DU TYPE STRUCTURAL	68
2.7 LE CODAGE ET LE CONTRE-CODAGE	72
2.8 L'ANALYSE DES DONNÉES ET LES CONCLUSIONS.....	72
3. OBJECTIF 2 : DESCRIPTION DES STRATÉGIES DE RÉOLUTION DE PROBLÈMES LEXICAUX.....	73
3.1 NOTRE CHOIX MÉTHODOLOGIQUE : LES ENTRETIENS MÉTAGRAPHIQUES	73
3.1.1 L'ENTRETIEN D'EXPLICITATION	73
3.1.2 LE DÉROULEMENT DE L'ENTRETIEN.....	74
3.1.3 LE FRUIT DE L'ENTRETIEN D'EXPLICITATION : LES VERBALISATIONS	75
3.1.4 LES PROBLÈMES INHÉRENTS À L'ENTRETIEN D'EXPLICITATION	76

3.1.5 LE RÔLE DE L'INTERVIEWEUR	77
3.1.6 L'ENTRETIEN MÉTAGRAPHIQUE.....	80
3.1.7 L'ENTRETIEN MÉTAGRAPHIQUE COMME OUTIL DE RECHERCHE.....	80
3.1.8 L'ENTRETIEN MÉTAGRAPHIQUE EN DYADE OU EN TRIADE	82
3.2 L'INSTRUMENT DE CUEILLETTE DES DONNÉES	83
3.3 LA CONSTITUTION DU CORPUS DE VERBALISATIONS.....	84
3.3.1 L'ÉCHANTILLON	84
3.3.2 L'ADMINISTRATION DES TESTS.....	85
3.3.3 LA TRANSCRIPTION DES ENTREVUES.....	86
3.4 LE CODAGE DES DONNÉES.....	86
3.4.1 LE CODAGE DES SÉQUENCES	87
3.4.2 LE CODAGE DES COMMENTAIRES.....	89
3.5 L'ANALYSE DES DONNÉES	95
3.5.1 L'ANALYSE DES SÉQUENCES	96
3.5.2 L'ANALYSE DES COMMENTAIRES	96
3.6 LIMITES DE LA RECHERCHE.....	96

IV. LES PROBLÈMES LEXICAUX DES ÉTUDIANTS UNIVERSITAIRES ... 98

1. PRÉSENTATION DES RÉSULTATS (OBJECTIF 1)	99
1.1 LE CODAGE ET LE CONTRE-CODAGE : RÉSULTATS	99
1.1.1 LE CONTRE-CODAGE : PREMIÈRE ANALYSE	100
1.1.2 LE CONTRE-CODAGE : DEUXIÈME ANALYSE	102
1.2 LES ERREURS LEXICALES : RÉSULTATS	103
1.3 LES MALADRESSES LEXICALES : RÉSULTATS.....	105
1.4 LES PROBLÈMES LEXICAUX : RÉSULTATS GLOBAUX	107
2. INTERPRÉTATION DES RÉSULTATS (OBJECTIF 1).....	109
2.1 LES PROBLÈMES LEXICAUX : PORTRAIT GLOBAL	109
2.2 LES PROBLÈMES LEXICAUX RELATIFS À LA COMBINATOIRE	110
2.2.1 LA COMBINATOIRE LEXICALE.....	111
2.2.2 LE RÉGIME	112
2.2.3 LE CARACTÈRE FIGÉ DES LOCUTIONS	113
2.2.4 LES PROPRIÉTÉS INHÉRENTES À L'UNITÉ LEXICALE	113
2.2.5 LE MODE VERBAL IMPOSÉ PAR L'UNITÉ LEXICALE.....	114
2.2.6 L'UTILISATION D'UN MAUVAIS VERBE AUXILIAIRE	114
2.3 LES PROBLÈMES LEXICAUX RELATIFS À LA FORME.....	115
2.3.1 LES ERREURS D'ORTHOGRAPHE.....	116
2.3.2 LES BARBARISMES.....	117
2.4 LES PROBLÈMES LEXICAUX DÉCOULANT D'UN USAGE DOUTEUX PAR RAPPORT AU CONTEXTE.....	118
2.4.1 LES PROBLÈMES DE REGISTRES DE LANGUE	118
2.4.2 LES RÉPÉTITIONS	119
2.4.3 LES MÉLANGES DE CODES	120
2.4.4 LES PLÉONASMES	120
2.4.5 LES CAS D'IMPRÉCISION.....	120
2.5 LES CAS D'UTILISATION D'UNE LEXIE FICTIVE ET LES PROBLÈMES DE SENS	121
2.5.1 LES IMPROPRIÉTÉS ET LES AUTRES PROBLÈMES DE SENS.....	122
2.5.2 LES CAS DE CRÉATION D'UNE NOUVELLE ACCEPTION D'UN VOCABLE	126

2.5.3 LES CAS DE MAUVAISES DÉRIVATIONS MORPHOLOGIQUES.....	127
2.6 LES CAS DE DOUBLE CODAGE.....	128
2.7 LES ERREURS D'INATTENTION.....	130
2.8 LE CAS LITIGIEUX.....	130
2.9 SUGGESTIONS RELATIVES À LA TYPOLOGIE DES PROBLÈMES LEXICAUX.....	131
2.9.1 LES PROBLÈMES D'ALTÉRATION DU SENS ET DES PROPRIÉTÉS DE COMBINATOIRE GRAMMATICALE	133
2.9.2 LES CAS D'UTILISATION D'UNE LEXIE FICTIVE.....	133
2.9.3 LES PROPRIÉTÉS DE COMBINATOIRE LEXICALE.....	134
CONCLUSION	136
 V. LES STRATÉGIES DE RÉOLUTION DE PROBLÈMES LEXICAUX DES ÉTUDIANTS UNIVERSITAIRES	 138
1. PRÉSENTATION DES RÉSULTATS (OBJECTIF 2)	139
1.1 LES SÉQUENCES : RÉSULTATS	139
1.1.1 PARTIE 1 : REPÉRAGE D'ERREURS LEXICALES	139
1.1.2 LA RÉOLUTION DES PROBLÈMES LEXICAUX	141
1.2 LES COMMENTAIRES : RÉSULTATS.....	144
1.2.1 LES COMMENTAIRES : PORTRAIT GÉNÉRAL	145
1.2.2 LES TYPES DE COMMENTAIRES	147
1.2.3 LES STRATÉGIES EN FONCTION DES TYPES DE PROBLÈMES LEXICAUX.....	149
2. INTERPRÉTATION DES RÉSULTATS (OBJECTIF 2).....	152
2.1 LES SÉQUENCES	152
2.1.1 LE REPÉRAGE DES PROBLÈMES LEXICAUX CIBLES	152
2.1.2 LES PROBLÈMES LEXICAUX NON CIBLES	156
2.1.3 LA RÉOLUTION DES PROBLÈMES LEXICAUX DANS LES DEUX PARTIES DU TEST.....	158
2.1.4 LA RÉOLUTION DES PROBLÈMES LEXICAUX PAR TYPES DE PROBLÈMES.....	159
2.2 LES COMMENTAIRES.....	165
2.2.1 LES RÉSULTATS GLOBAUX	165
2.2.2 LES TYPES DE COMMENTAIRES	169
2.2.3 LES STRATÉGIES DÉPLOYÉES EN FONCTION DES TYPES DE PROBLÈMES.....	174
CONCLUSION	183
 VI. CONCLUSION.....	184
1. RAPPEL DES PRINCIPAUX RÉSULTATS	185
2. LES RECHERCHES FUTURES	186
3. LINGUISTIQUE ET DIDACTIQUE.....	188
4. LES CONCLUSIONS.....	189
4.1 LES PISTES LINGUISTIQUES.....	190
4.2 LES PISTES DIDACTIQUES	191
 BIBLIOGRAPHIE	xiv
 ANNEXE I : formulaires de consentement.....	xxi
ANNEXE II : test du deuxième objectif.....	xxiv
ANNEXE III : consignes pour le test du deuxième objectif	xxix

LISTE DES TABLEAUX

CHAPITRE II

Tableau I. <i>Étapes de la résolution d'un problème</i>	46
Tableau II. <i>Étapes de la résolution d'un problème lexical</i>	50

CHAPITRE III

Tableau III. <i>Typologie structurale des problèmes lexicaux</i>	65
Tableau IV. <i>Grille de codage des séquences</i>	87
Tableau V. <i>Grille de codage des commentaires</i>	91

CHAPITRE IV

Tableau VI. <i>Types d'erreurs lexicales rencontrées</i>	103
Tableau VII. <i>Types de maladresses lexicales rencontrées</i>	105
Tableau VIII. <i>Types de problèmes lexicaux rencontrés</i>	107
Tableau IX. <i>Typologie structurale des problèmes lexicaux (révisée)</i>	132

CHAPITRE V

Tableau X. <i>Problèmes cibles repérés dans la première partie du test</i>	140
Tableau XI. <i>Problèmes cibles bien et mal résolus et statuts des séquences correspondantes</i>	143
Tableau XII. <i>Statuts des commentaires en fonction de l'étape de résolution du problème</i>	145
Tableau XIII. <i>Données concernant les 1531 commentaires codés</i>	146
Tableau XIV. <i>Types de commentaires : nombre et pourcentage des 1531 commentaires codés</i>	148
Tableau XV. <i>Stratégies les plus fréquentes selon les types de problèmes lexicaux..</i>	151

LISTE DES FIGURES

CHAPITRE IV

Figure 1. <i>Proportions des différents types d'erreurs rencontrées</i>	104
Figure 2. <i>Proportions des différents types de maladresses rencontrées</i>	106
Figure 3. <i>Proportions des différents types de problèmes lexicaux rencontrés</i>	108
Figure 4. <i>Les problèmes de combinatoire</i>	111
Figure 5. <i>Les problèmes de forme</i>	115
Figure 6. <i>Les problèmes découlant d'un usage douteux en fonction du contexte</i>	118
Figure 7. <i>Les cas d'utilisation d'une lexie fictive et les problèmes de sens</i>	121

LISTE DES CONVENTIONS D'ÉCRITURE

Signifié (sens linguistique)	: 'livre'
Signifiant (mot-forme)	: <i>guichet automatique</i>
Unité lexicale (lexie)	: RÉFRIGÉRATEUR
Acceptions d'un vocable	: CONSTRUCTION ₁ , CONSTRUCTION ₂
Vocables homonymes	: GRÈVE ¹ , GRÈVE ²
Agrammaticalité	: * <i>le tortue</i>
Asémantisme	: # <i>La table sourit.</i>
Acceptabilité douteuse	: ? <i>un vaste chapeau</i>
Préposition régie	: ENCLIN (à)
Extrait de texte d'étudiant	: ... les méthodes de punition et de félécitation ... [sujet K-14]
Extrait d'entrevue	: Pasque... ça s'dit pas ! [équipe 5, séq. 5]
Numéro de chapitre	: I, II, III...
Numéro de section	: II-3.1.2

*À mes parents,
pour leur appui inconditionnel.*

REMERCIEMENTS

La rédaction d'un premier travail de recherche d'envergure demande un encadrement constant et implique la participation de plusieurs personnes. J'aimerais donc exprimer ma reconnaissance à ces gens qui m'ont permis, par leurs conseils et leur générosité, de mener à terme ce mémoire.

J'aimerais d'entrée de jeu témoigner ma profonde gratitude à Pascale Lefrançois, qui a su me guider dans mon apprentissage de la recherche et m'initier au monde de la didactique. Par ses innombrables relectures et ses multiples suggestions, elle a grandement contribué à la cohérence et à la pertinence du présent mémoire. Je désire saluer sa grande disponibilité, son attention méticuleuse et vigilante, mais aussi sa bonne humeur et son sens de l'humour qui ont été une source perpétuelle d'encouragement tout au long de ces deux années.

Le volet linguistique de ce mémoire doit beaucoup aux remarques éclairées d'Alain Polguère. Il a su, par ses commentaires critiques, son implacable recherche de cohésion et son regard lucide sur le fonctionnement de la langue, contribuer grandement à la validité linguistique de ce travail. Le côtoyer au cours des dernières années m'a permis d'acquérir, au-delà des connaissances linguistiques et des expériences de travail lexicographique, une fascination pour les phénomènes de sens et une passion pour le lexique, cette composante essentielle de la langue qui se plie pourtant difficilement à la description. De plus, son écriture fluide et limpide et son style concis constituent de grandes sources d'inspiration.

Je voudrais aussi exprimer ma reconnaissance à Igor Mel'čuk qui, bien qu'il n'ait pas participé directement au présent mémoire, alimente sans cesse mes réflexions linguistiques. Il m'a permis, en m'offrant le privilège de travailler à ses côtés, de mieux saisir la pertinence, la cohérence et l'élégance de la théorie Sens-Texte, de sorte qu'il me semble aujourd'hui essentiel d'en effectuer une transposition didactique afin que le monde de l'enseignement puisse en bénéficier.

Je voudrais aussi remercier les enseignantes du cours *Français écrit pour futurs enseignants* qui m'ont aidé, à l'automne 2004, dans la collecte de mon corpus de textes en m'ouvrant les portes de leur classe. Elles m'ont aussi permis de me servir de leurs corrections des copies d'étudiants pour la vérification de ma propre correction. Je tiens donc à souligner l'aide précieuse de Monique Gervais, Danielle Lefebvre et Denise Sabourin.

Évidemment, la constitution de mon corpus d'analyse aurait été impossible sans la généreuse participation de quelque 120 étudiants. Je remercie donc les étudiants du cours *Français écrit pour futurs enseignants* qui ont accepté que leur rédaction soit utilisée dans le cadre de ce mémoire. Je tiens aussi à exprimer ma profonde reconnaissance aux 16 participants des entrevues menées au printemps 2005, qui m'ont gracieusement accordé quelques heures de leur temps dans cette tumultueuse période d'après-grève.

Merci aussi à ma collègue et amie Ophélie Tremblay, qui a assuré le contre-codage d'une partie de mon corpus d'erreurs. Elle a contribué, par ces judicieuses remarques, à l'amélioration de ma typologie des problèmes lexicaux.

Finalement, je désire souligner l'appui du Conseil de recherche en sciences humaines du Canada, qui m'a octroyé une bourse en 2004-2005, ainsi que le soutien de la Faculté des sciences de l'éducation qui m'a attribué, en collaboration avec la Faculté des études supérieures, une bourse pour la rédaction de ce mémoire.

INTRODUCTION

La présente recherche, menée dans le cadre d'une maîtrise en didactique faite en codirection avec le département de linguistique et de traduction, s'intéresse à la maîtrise du lexique chez les étudiants universitaires. Elle a pour objectif de répondre à deux grandes questions :

- 1) Quels sont les problèmes lexicaux des étudiants universitaires?
- 2) Quelles sont les stratégies déployées par les étudiants universitaires dans la résolution de problèmes lexicaux?

Les réponses à ces questions devraient permettre de tirer certaines conclusions didactiques relatives à l'enseignement du lexique.

La structure de notre travail suit la forme classique des recherches en didactique. Le premier chapitre présente notre problématique de recherche. Le chapitre II constitue le cadre conceptuel du mémoire et expose les notions auxquelles nous ferons appel tout au long de la recherche. Le troisième chapitre explique les méthodologies relatives à nos deux objectifs de recherche et présente le principal apport théorique de notre travail : la typologie structurale des problèmes lexicaux. Les chapitres IV et V sont les seuls qui dérogent à la forme classique des rapports de recherches; comme nous avons deux objectifs de recherche distincts auxquels correspondent des résultats de natures différentes, nous avons choisi de ne pas adopter la formule classique du chapitre de présentation des résultats suivi du chapitre d'interprétation. Ainsi, les chapitres IV et V comportent tous les deux un volet de présentation des résultats et un volet d'interprétation. Étant donné la quantité d'informations présentées pour chacun des objectifs, nous croyons qu'une telle division des chapitres favorisera la compréhension et la cohésion du mémoire. Finalement, le chapitre VI tirera des conclusions linguistiques et didactiques des résultats analysés dans les chapitres précédents.

Nous espérons que cette recherche saura éclairer un peu la question de la maîtrise du lexique par les étudiants, zone de la maîtrise du code écrit peu explorée jusqu'à maintenant. Nous souhaitons aussi qu'elle ouvrira la voie à d'autres recherches qui délaisseront cette fois la perspective descriptive pour proposer des solutions aux problèmes constatés ici.

I. PROBLÉMATIQUE

1. CONTEXTE DE LA RECHERCHE

L'état du français préoccupe les Québécois depuis plusieurs années déjà, notamment en ce qui concerne la maîtrise du code écrit chez les étudiants. Une étude menée en 1995 démontre que « plus du tiers des étudiants des ordres postsecondaires ne disposent pas d'une maîtrise suffisante du français écrit, en particulier en ce qui concerne les relations grammaticales et l'orthographe lexicale de base » (Roy, Lafontaine et Legros, 1995 : 30). Cette situation est d'autant plus alarmante qu'elle met en péril la réussite des étudiants aux ordres supérieurs d'enseignement.

Sensibles à ce problème, le ministère de l'Éducation du Québec (MÉQ) ainsi que les différents établissements d'enseignement ont instauré divers tests dans le but d'évaluer les compétences linguistiques des étudiants et de s'assurer qu'un certain niveau de maîtrise du français écrit est atteint. Ainsi, le MÉQ considère maintenant la maîtrise du code linguistique comme une condition à l'obtention du diplôme d'études secondaires et exige la réussite de l'épreuve uniforme de français en fin d'études collégiales (Conseil de la langue française, 1998). Certaines universités évaluent aussi la qualité du français à l'admission par divers tests comme le *Test universitaire de rédaction et de bonne orthographe (TURBO)* de l'Université de Sherbrooke et le *Test de français Laval-Montréal (TFLM)* de l'Université Laval et de l'Université de Montréal. Des mesures d'aide ont aussi été mises sur pied dans les cégeps et les universités du Québec, comme les Centres d'aide en français (CAF) et les cours de mise à niveau. Nous constatons que beaucoup d'efforts sont faits dans le but d'améliorer la situation (Table Éducation Montréal, 2002).

Pour bien comprendre les problèmes de maîtrise du code linguistique chez les élèves québécois, il faut nécessairement se demander quelles connaissances et quelles compétences leur font défaut. Évidemment, beaucoup de travaux ont tenté de répondre à cette question en recensant les différents types d'erreurs commises par les élèves (Bureau, 1985 ; ministère de l'Éducation, 1997 ; groupe DIEPE, s.d.), mais peu, à notre connaissance, se sont intéressés aux processus qui mènent aux erreurs. Pour intervenir efficacement auprès des élèves, il faut non seulement savoir quels sont les types d'erreurs commises, mais il faut aussi tenter de comprendre ce qui amène les étudiants à les commettre. De plus, très peu de travaux se sont penchés sur la maîtrise du français écrit aux ordres supérieurs d'enseignement. Pourtant, si c'est à ce niveau que le code linguistique cesse d'être étudié comme objet propre, sa maîtrise continue d'être – et plus que jamais – une compétence « transversale » et fondamentale à

laquelle l'étudiant doit sans cesse faire appel ; la maîtrise du français écrit est désormais tenue pour acquise, et considérée comme une condition *sine qua non* à la poursuite d'études supérieures.

Notons quand même deux recherches, menées au Québec, qui ont porté sur la maîtrise du français écrit chez les étudiants du postsecondaire. Dans leur ouvrage *Le Savoir grammatical après treize ans de formation*, Roy et al. (1995) ont tenté d'éclairer cette question par l'analyse de verbalisations faites par 165 étudiants québécois du collégial et de l'université confrontés à des activités de révision et de production de textes. Cette étude mène à deux conclusions principales. D'abord, « quinze ans de formation grammaticale dont les assises reposent surtout sur la mémoire et sur l'usage de trucs [conduit les étudiants à] peu de cohérence [et à] beaucoup de psittacisme » (Roy, Boudreau, Viau et Lafontaine, 1993). En d'autres mots, bien que les étudiants se souviennent généralement des règles et « trucs » grammaticaux appris au cours de leur formation, ils n'arrivent que très rarement à les appliquer de façon efficace, c'est-à-dire en fonction du contexte du problème à résoudre. Ensuite, les étudiants ne maîtrisent pas le métalangage qui leur permettrait de verbaliser et de réfléchir sur la langue. Comme l'étude de la langue est particulière en ce sens que son objet s'utilise lui-même pour se décrire, le métalangage est un outil précieux favorisant une réflexion efficace et cohérente qui permet de différencier l'objet d'étude de l'étude elle-même. C'est cette capacité de recul qui semble absente chez beaucoup d'étudiants universitaires et qui est, entre autres, à la source d'une maîtrise plutôt précaire du code écrit. Une autre étude (Lefrançois, 2004), menée auprès d'étudiants du 1^{er} cycle universitaire, arrive à des résultats similaires, à savoir que les étudiants utilisent peu le métalangage et qu'ils parviennent difficilement à adapter leurs stratégies de résolution de problèmes au contexte. Cette étude a de plus démontré que les étudiants ont recours à des stratégies insuffisantes et souvent incorrectes.

2. PROBLÈME GÉNÉRAL

Il ressort aussi de ces deux études que les erreurs les moins repérées et les moins bien résolues sont celles de syntaxe, de lexique et de cohérence textuelle. Par contre, les deux recherches s'étant plutôt intéressées aux erreurs d'orthographe grammaticale, relativement peu de données ont été recueillies en ce qui concerne les erreurs relevant

de ces dimensions. Conséquemment, nous détenons peu d'informations sur les stratégies de résolution de problèmes s'y rapportant.

3. ANALYSE DU PROBLÈME

Des trois dimensions laissées dans l'ombre par les précédents travaux, à savoir la syntaxe, la cohérence textuelle et le lexique, nous avons choisi de nous intéresser principalement à la dernière. Notre choix est bien sûr motivé par un intérêt personnel pour les phénomènes lexicaux, mais est aussi appuyé par des motifs plus profonds. En effet, nous croyons qu'il est impératif de se pencher sur cette dimension de la maîtrise de la langue qui n'a, jusqu'à maintenant, été que peu documentée dans une perspective didactique. Énormément de travail est fait par les linguistes (lexicologues et sémanticiens) en ce qui concerne le fonctionnement et la structure du lexique, et de nombreux phénomènes lexicaux sont maintenant très bien décrits, mais, à notre connaissance, rares sont les ponts établis entre ces concepts et la didactique du français, qui pourrait pourtant en tirer profit. C'est un peu dans cette perspective d'« application » des concepts linguistiques à la didactique que nous choisissons de nous intéresser particulièrement aux problèmes lexicaux des étudiants. Et parce que nous croyons que le lexique, en tant que véhicule des idées et composante essentielle de la langue, devrait prendre une place plus importante dans la compréhension du code linguistique et dans l'étude de son fonctionnement.

3.1 LE LEXIQUE AU POSTSECONDAIRE

En consultant les travaux sur la maîtrise du français écrit au postsecondaire, on s'aperçoit que le lexique constitue la dimension la moins prise en compte par les étudiants (Roy *et al.*, 1995). La recherche de Lefrançois (2004) démontre aussi que les erreurs de lexique sont parmi les moins repérées et les moins souvent résolues correctement. Ceci dénote un manque flagrant d'outils permettant aux étudiants d'aborder de façon efficace une erreur de lexique. Paradoxalement, si l'on se fie aux résultats de l'épreuve uniforme du MÉQ (résultats de 2003-2004), 90,4 % des étudiants ont commis 2 fautes ou moins de vocabulaire (MÉQ, 2004). Lépine (1995 : 24-30) nous invite pourtant à nous méfier de ces résultats encourageants : les tests

traditionnels visant à mesurer la maîtrise du code linguistique, comme la production de textes, s'avèrent beaucoup moins efficaces pour évaluer la « maîtrise¹ lexicale » d'un individu. En effet, « on s'est plutôt rendu compte qu'un texte de 500 mots ne constituait pas un échantillon suffisant pour vérifier la maîtrise du lexique chez un individu – sans parler de la difficulté de cerner et de noter la faute de vocabulaire. » (Lépine, 1995 : 24).

Par ailleurs, en observant les programmes de français du MÉQ, force est de constater que « l'enseignement du lexique apparaît comme un parent pauvre en didactique du français langue maternelle » (Simard, 1994 : 28). Même ailleurs dans la francophonie, le lexique semble donner lieu à un enseignement plutôt sporadique. Comme le souligne Dreyfus (2004), les enseignants du primaire en France dénoncent les difficultés lexicales de leurs élèves, mais avouent par ailleurs ne consacrer que très peu de temps à cette dimension dans leur enseignement. Au Québec, des enseignants du collégial et de l'université interviewés dans le cadre de la recherche de Lefrançois, Laurier, Lazure et Claing (2005) considèrent que leurs étudiants ont des problèmes de vocabulaire. Par contre, même s'ils ont en début d'année l'intention de travailler le vocabulaire, ils perçoivent en fin d'année peu d'impact de leur enseignement sur la maîtrise lexicale des étudiants.

De plus, lorsqu'on s'intéresse au lexique, il est souvent approché selon une perspective plutôt superficielle se résumant la plupart du temps à l'orthographe lexicale, certainement sa dimension la plus accessoire. Simard (1994) relève d'ailleurs l'emploi abusif du terme *liste de vocabulaire* très répandu en didactique du français, et qui renvoie en fait à des listes d'orthographe lexicale. Cette confusion de termes en apparence banale dénote une assimilation du concept de vocabulaire à la simple dimension orthographique. Fortier (1994) constate, pour sa part, que l'enseignement direct du vocabulaire² constitue à la fois la méthode la plus répandue et la moins efficace.

La question du lexique au postsecondaire soulève donc un triple problème. Premièrement, les étudiants semblent dépourvus face à un problème lexical. Deuxièmement, les tests auxquels nous soumettons les étudiants ne parviennent pas à évaluer le niveau de maîtrise lexicale de façon efficace et ne permettent par

¹ Nous définirons ce que nous entendons par « maîtrise » à la section I-3.2.

² L'enseignement direct du vocabulaire consiste à présenter des unités lexicales difficiles ou techniques aux élèves et à leur demander d'en trouver les définitions. Comme les élèves ne savent pas toujours utiliser efficacement les ouvrages de référence, ils tirent souvent d'un tel enseignement des connaissances lacunaires ou erronées.

conséquent pas de saisir l'incidence du lexique sur la maîtrise « générale » du code linguistique. Et finalement, les formations primaire et secondaire ne semblent pas fournir aux élèves désireux d'entreprendre des études aux cycles supérieurs d'enseignement les outils nécessaires pour favoriser leur compétence lexicale.

3.2 LA MAÎTRISE ET LA COMPÉTENCE LEXICALES

Il semble important de faire ici une distinction entre les termes *maîtrise* et *compétence*. Nous considérerons, dans le cadre de cette recherche, que la *maîtrise* est plutôt liée à la connaissance d'unités lexicales³ particulières. Ainsi, nous pourrions dire que la *maîtrise lexicale* est locale, puisque propre à chaque unité lexicale. Il sera donc impossible d'évaluer la maîtrise lexicale d'une personne sans se baser sur la maîtrise qu'il a d'un nombre plus ou moins restreint d'unités lexicales (par exemple, celles qu'il utilise dans un texte ou encore celles qui lui sont soumises à l'occasion d'un test de vocabulaire). Nous verrons dans notre cadre conceptuel que l'on évitera de dire qu'un étudiant « maîtrise bien le lexique du français », mais plutôt qu'il « maîtrise bien un certain vocabulaire ». La *compétence lexicale*, pour sa part, serait l'ensemble des connaissances et stratégies permettant de comprendre les phénomènes lexicaux et d'acquérir ainsi plus efficacement la maîtrise de nouvelles unités lexicales. Elle serait donc un accès à la maîtrise, bien que cette dernière puisse tout à fait exister sans la compétence correspondante.

3.3 LE LEXIQUE AU-DELÀ DE L'ORTHOGRAPHE LEXICALE

L'orthographe d'usage est souvent considérée comme le problème majeur concernant le lexique ; mais un « mot » n'est-il pas plus qu'une association de symboles? Bien entendu, l'orthographe lexicale ou forme écrite d'une unité lexicale (ce que Saussure appelait le signifiant du signe linguistique, voir II-1.3.1) est primordiale puisqu'elle en est l'accès ; c'est sans doute sa dimension la plus tangible. Mais au-delà de ce contenant, qu'en est-il du contenu d'une unité lexicale ? Est-ce que la connaissance d'une unité lexicale se résume à la maîtrise de sa forme écrite ? Bien sûr que non.

³ Pour des raisons qui seront exposées dans notre cadre conceptuel, nous éviterons d'employer le terme *mot*, qui dénote plusieurs types d'entités linguistiques. Pour une définition d'*unité lexicale*, ou *lexie*, voir section II-1.3.4.

Schmitt (2000) propose une liste des différentes connaissances nécessaires pour considérer complète la maîtrise d'une unité lexicale (traduction libre) :

- le sens du mot (polysémie, homonymie)
- la forme écrite du mot (orthographe d'usage)
- la forme parlée du mot
- le comportement grammatical du mot (partie du discours, régime, prépositions régies)
- les collocations contrôlées par le mot (collocation, locution)
- le registre de langue auquel le mot appartient et sa fréquence
- les relations que le mot entretient avec d'autres mots (champ lexical)

« These are known as types of word knowledge, and most or all of them are necessary to be able to use a word in the wide variety of language situations one comes across. » (Schmitt, 2000 : 5). En effet, plusieurs dimensions doivent être considérées à la fois lors de l'utilisation d'une unité lexicale ; par conséquent, les occasions d'erreurs sont nombreuses lorsqu'on manie le lexique. Ainsi nous est-il permis de supposer que d'une connaissance défaillante des phénomènes lexicaux découleront plusieurs types d'erreurs.

L'erreur de lexique demeure une notion floue. Certains problèmes souvent attribués à la syntaxe ont en fait leur source dans le lexique. Par exemple, Roy (1995) définit l'usage de la préposition et le régime verbal comme des contenus syntaxiques. À notre avis, ces deux sources de problèmes chez nombre d'étudiants au postsecondaire ne relèvent pas de la syntaxe, du moins pas de la syntaxe de la phrase, mais plutôt du lexique, en ce sens que les prépositions régies et les régimes verbaux sont des propriétés inhérentes à chaque unité lexicale. Cette microsyntaxe régie par l'unité lexicale, que Lépine (1995) propose d'appeler « syntaxe lexicale », relève donc en effet du lexique, puisque c'est par elle que chaque unité lexicale imposera ses caprices à la grammaire du français. La connaissance d'une unité lexicale implique donc aussi la connaissance des compléments qu'elle permet ou exige et des prépositions qui servent à les introduire. Ainsi, *on remédie à qqch.* mais *on pallie qqch.*, *on monte au grenier* mais *en chaire*. Roy (1995) attribue aussi à la syntaxe d'autres erreurs faites par les étudiants du postsecondaire, comme des maladresses dans l'utilisation des connecteurs. Encore une fois, nous pourrions avancer que de telles erreurs sont plutôt d'ordre lexical, puisqu'elles traduisent un manque de maîtrise du sens de ces connecteurs. En fait, le connecteur sera la plupart du temps « bien placé »

syntactiquement, mais c'est de son sens, un aspect lexical, que découlera une erreur logique dans ce contexte. Comme le mentionne Denance (1999) :

« Un des défis que nous avons en tant que correcteurs et superviseurs, c'est de faire le départage entre les fautes de vocabulaire et celles de syntaxe. [...] La frontière entre syntaxe et sémantique est en effet poreuse : il faut souvent retourner au sens pour détecter et corriger une faute de syntaxe... Si on joue sur le sens, ne sommes-nous pas alors devant une faute de vocabulaire? Beaucoup de fautes hâtivement pénalisées en syntaxe se révèlent en fait de « simples » fautes de vocabulaire. »

Nous croyons donc que l'erreur lexicale a une portée beaucoup plus large que celle qui lui est généralement admise. Le lexique contenant l'ensemble des informations et caractéristiques inhérentes à chaque unité lexicale, les répercussions de cet aspect de la langue sur les autres dimensions du code sont plus importantes que ce que l'on considère habituellement. Le premier pas vers une compréhension de la maîtrise lexicale des étudiants consistera donc nécessairement à caractériser clairement la notion d'erreur lexicale.

4. PROBLÈME SPÉCIFIQUE

En bref, nous observons qu'il règne une certaine confusion autour de la notion d'erreur lexicale. Il ressort aussi des deux études mentionnées plus haut (Roy *et al.*, 1995 ; Lefrançois, 2004) que le lexique constitue l'une des dimensions les moins considérées par les universitaires du 1^{er} cycle lors d'une activité de révision ou de rédaction de texte. Par contre, si ces deux recherches mentionnent que le lexique s'avère l'une des principales lacunes, elles ne précisent pas quels types d'erreurs sont commises à l'intérieur de la dimension lexicale. De plus, il subsiste un manque de connaissances quant aux stratégies de résolution de problèmes lexicaux des étudiants selon les types d'erreurs commises.

5. OBJECTIFS DE RECHERCHE

Voici donc les objectifs visés par la présente recherche :

1. décrire les types de problèmes lexicaux d'étudiants universitaires soumis à une activité de rédaction de texte et établir la fréquence de chaque type ;
2. décrire les stratégies de résolution de problèmes lexicaux d'étudiants universitaires, notamment en fonction des types de problèmes rencontrés.

6. PERTINENCE ET VISÉES DE LA RECHERCHE

Pour être en mesure d'offrir aux étudiants des mesures d'aide adaptées à leurs problèmes, il est primordial de connaître quels aspects du lexique occasionnent des erreurs. Mais au-delà de cette information « de surface », nous devons aussi comprendre de quelle façon ils réfléchissent lorsqu'ils sont confrontés à un problème de nature lexicale. Cette recherche vise donc à éclairer une dimension de la compétence écrite qui demeure obscure même après les quelques études s'intéressant aux stratégies de résolution de problèmes en français écrit. De plus, elle a comme sujets des étudiants universitaires, population que peu de travaux ont étudiée en didactique du français. En mettant en lumière l'incidence de l'erreur lexicale sur la maîtrise du code écrit, nous espérons faire ressortir l'importance d'une « compétence lexicale », et permettre ainsi de revoir les méthodes d'enseignement du lexique ou du vocabulaire.

II. CADRE CONCEPTUEL

Le cadre conceptuel constitue la charpente théorique de tout projet de recherche, et doit par conséquent présenter l'ensemble des notions auxquelles fera appel la méthodologie. La présentation de notre cadre conceptuel se divise en quatre parties. La première, plus proche du cadre théorique, a pour mission de définir clairement l'ensemble des termes relatifs au lexique que nous utiliserons dans la présente recherche. La deuxième partie, quant à elle, vise à caractériser la notion d'erreur lexicale essentielle à la réalisation de notre premier objectif. La troisième section s'intéresse aux concepts de problème et de stratégies de résolution de problèmes en général, et tente de déterminer de quelle façon ces notions peuvent être transposées au lexique. Finalement, la quatrième partie présentera le concept de métalangage auquel nous ferons appel dans la réalisation de notre deuxième objectif de recherche.

1. LE LEXIQUE

La présente section vise à poser les assises théoriques qui nous permettront ensuite de tenir un discours cohérent sur notre objet de recherche : le lexique. Nous situerons d'abord notre cadre théorique en présentant succinctement la théorie Sens-Texte et la Lexicologie explicative et combinatoire (LEC). Nous introduirons ensuite certaines notions essentielles à une réflexion sur la langue avant d'expliquer pourquoi nous éviterons d'employer le terme *mot*, pourtant intimement lié au concept de lexique. À partir de cette explication, nous déploierons une terminologie qui nous permettra d'aborder la question de l'erreur de lexique avec clarté et précision. La section II-1.4 exposera les éléments constitutifs d'une définition lexicographique, qui nous permet d'explicitier le sens des unités lexicales. Finalement, la section II-1.5 s'intéressera à une dimension peu connue de l'unité lexicale : ses propriétés de combinatoire.

1.1 NOTRE CADRE THÉORIQUE : LA THÉORIE SENS-TEXTE

La théorie linguistique qui servira de toile de fond à l'élaboration de notre mémoire de maîtrise est la théorie Sens-Texte. Essentiellement développée par Igor Mel'čuk, professeur au département de linguistique et de traduction de l'Université de Montréal, cette théorie, qui a émergé à Moscou au début des années 60, est issue des courants structuraliste et fonctionnaliste. Elle conçoit la langue comme un système permettant

de faire correspondre à un sens donné l'ensemble des « textes » (au sens le plus large du terme) qui le véhiculent. Considérant le sens comme point de départ, la théorie Sens-Texte place la modélisation du lexique au centre de la modélisation de la langue. Elle modélise la langue en quatre niveaux principaux – sémantique, syntaxique, morphologique, phonologique (qui sont d'ailleurs les quatre domaines « traditionnels » de la linguistique) –, qui peuvent se diviser à leur tour. Elle envisage la langue selon une perspective synchronique (comme le propose Saussure, 1972) et, jusqu'à maintenant, s'est surtout intéressée aux niveaux sémantique, syntaxique et morphologique. La syntaxe Sens-Texte est une syntaxe de dépendance fortement inspirée de celle de Lucien Tesnière (1959), un des précurseurs du structuralisme en linguistique. Dans le cadre de notre recherche, nous ferons surtout appel à la Lexicologie Explicative et Combinatoire (LEC), branche lexicologique de la théorie Sens-Texte. La LEC permet de modéliser l'ensemble des phénomènes lexicaux à l'aide d'une approche formelle. En plus de décrire la forme (signifiant) et le sens (signifié) des unités lexicales, elle s'intéresse à tous les phénomènes de combinatoire lexicale (voir II-1.5), qui sont généralement décrits de façon informelle dans les dictionnaires commerciaux et presque systématiquement laissés pour compte dans l'enseignement du lexique. C'est là tout l'intérêt de cette théorie pour nous, puisqu'elle nous permettra de décrire avec finesse et exactitude les erreurs lexicales commises par les étudiants.

1.1.1 LA LANGUE SELON LA THÉORIE SENS-TEXTE

La théorie Sens-Texte propose une définition fonctionnelle de la langue (Mel'čuk, 1997 : 6).

La langue est considérée comme un mécanisme, ou un système de règles, qui permet au locuteur de faire deux choses :

1. *PARLER, c'est-à-dire (être capable de) faire correspondre à un sens qu'il veut exprimer tous les textes de sa langue qui, d'après lui, peuvent véhiculer ce sens et choisir celui qui passe le mieux dans les circonstances concrètes d'un acte langagier donné ;*
2. *COMPRENDRE LA PAROLE, c'est-à-dire (être capable de) faire correspondre à un texte qu'il perçoit tous les sens que, d'après lui, ce texte peut véhiculer et choisir celui qui passe le mieux dans les circonstances concrètes d'un acte langagier donné.*

Nous considérerons donc la langue comme un système permettant de traduire une pensée (sens) en parole (texte), dans le but de pouvoir communiquer cette pensée.

Ainsi pouvons-nous conclure que c'est le lexique qui véhicule vraiment le sens ; si la grammaire est la mécanique permettant la construction d'un message, le lexique en est le matériau, et c'est par lui que nous « touchons » à l'essence même du message. C'est là un postulat central de la théorie Sens-Texte qui stipule que la langue doit être décrite du sens vers le texte (Mel'čuk, 1997 : 14-15).

1.1.2 LA LEXICOLOGIE EXPLICATIVE ET COMBINATOIRE

La lexicologie explicative et combinatoire (LEC) constitue le volet lexicographique de la théorie Sens-Texte, celui qui nous intéressera particulièrement. L'originalité de la LEC réside dans le fait qu'elle s'intéresse, au-delà de la forme et du sens des unités lexicales, à leurs relations avec les autres unités lexicales de la langue. Par le formalisme des fonctions lexicales (voir Mel'čuk, Clas et Polguère, 1995), la LEC permet d'encoder l'ensemble des relations privilégiées qu'entretiennent les unités lexicales entre elles.

L'aboutissement ultime de la LEC est un type particulier de dictionnaire appelé *Dictionnaire explicatif et combinatoire (DEC)*. Certains volumes d'un dictionnaire de ce type ont été déjà publiés sous le nom de *Dictionnaire explicatif et combinatoire de français contemporain (DECFC)* et de *Explanatory Combinatorial Dictionary of Modern Russian*. Un DEC contient, en plus des informations traditionnelles à propos de la forme, du sens et des marques d'usage, l'ensemble des informations de combinatoire qui devraient permettre de « préciser les conditions nécessaires et suffisantes pour qu'[on] emploie le mot étudié et lui seul dans tous les cas où il doit être employé, autrement dit de « prédire » tous les emplois d'un mot » (Picoche, 1992 : 24). Comme l'ensemble du lexique français n'est pas encore décrit dans le DECFC, il ne pourra pas servir systématiquement d'ouvrage de référence (bien que la description de certains champs sémantiques soit plutôt exhaustive). Par contre, les principes de description lexicographique de la LEC constitueront un point de repère important en ce qui concerne les connaissances nécessaires à la maîtrise d'unités lexicales. Par conséquent, ils nous offriront des indices quant aux types d'erreurs qui peuvent être commises en maniant le lexique.

1.2 DEUX MISES AU POINT TERMINOLOGIQUES

Il importe de clarifier ici certains termes qui sont utilisés de diverses façons dans la vie quotidienne, notamment en enseignement, mais qui renvoient à des concepts spécifiques fondamentaux pour quiconque désire entreprendre une réflexion linguistique.

1.2.1 LA DICHOTOMIE *LANGUE* ~ *PAROLE*

Nous avons vu plus haut que la langue est un système permettant d'assembler les unités lexicales selon des règles bien précises. Polguère (2003 : 16) définit la langue comme « un système de signes conventionnels et de règles de combinaison de ces signes, qui forment un tout complexe et structuré ». Pour fonctionner, ce système doit bien entendu être commun à un ensemble d'individus, qui forment une communauté linguistique. La langue est donc un outil qui permet la communication entre des gens partageant un même système.

Cette communication se fait par la parole, qui est une actualisation de la langue. La langue se manifeste et s'utilise par la parole, et ne peut être observée qu'en observant la parole, « c'est-à-dire [...] les instances d'échanges langagiers entre au moins deux individus : le locuteur et le destinataire » (Polguère, 2003 : 19).

1.2.2 LA DICHOTOMIE *LEXIQUE* ~ *VOCABULAIRE*

Une distinction importante existe entre les notions de lexique et de vocabulaire, bien que ces termes soient souvent utilisés indifféremment l'un pour l'autre. La présente recherche s'intéressant au lexique, il importe ici de souligner la différence entre ces deux notions.

Le lexique est une composante de la langue : c'est l'« entité théorique qui correspond à l'ensemble des lexies de cette langue » (Polguère, 2003 : 70). Un locuteur ne pourra jamais maîtriser le lexique d'une langue dans sa totalité. On dit d'ailleurs que le lexique est une entité théorique, puisqu'il est impossible de recenser l'ensemble des unités lexicales d'une langue; les dictionnaires, par exemple, qui sont souvent perçus comme la référence en matière de lexique, en omettent beaucoup (régionalismes, termes techniques, néologismes, etc.).

Le vocabulaire, quant à lui, est un sous-ensemble du lexique. Le vocabulaire d'un individu est l'ensemble des unités du lexique qu'il connaît et maîtrise. Le vocabulaire d'un texte sera l'ensemble des unités lexicales que ce texte contient. Le vocabulaire étant un concept propre à un individu ou à un texte, il faudra éviter de parler du *vocabulaire d'une langue*.

1.3 LE TERME MOT : SOURCE DE CONFUSION

L'utilisation du terme *mot* est source de nombreuses confusions dans l'enseignement, puisque ce terme est utilisé pour désigner plusieurs concepts différents. Pour pouvoir tenir un discours efficace sur le lexique, la première étape consiste donc selon nous à désambiguïser le terme *mot* – que nous éviterons dorénavant – pour ensuite accoler un nom plus précis à chacun des concepts qu'il englobe. Seront donc introduits les termes *mot-forme*, *lexème*, *compositionnalité sémantique*, *locution ou expression idiomatique*, *expression linguistique*, *unité lexicale* ou *lexie*, *polysémie*, *vocable*, *acception*, *homonymie*, *synonymie* et *paraphrase*.

1.3.1 LES TROIS DIMENSIONS DU SIGNE LINGUISTIQUE

Le signe linguistique, tel que défini par Ferdinand de Saussure (1972), est une association entre un contenu (**signifié** ou **sens**) et une forme (**signifiant**). Chaque signe linguistique devra donc avoir un sens et un seul, auquel nous aurons accès par l'intermédiaire de son signifiant, qui pourra être une chaîne de phonèmes ou, dans le cas du code écrit, une chaîne de caractères (Saussure parlait plutôt d'*image acoustique*). Comme la présente recherche porte sur le lexique, nous nous intéresserons exclusivement aux signes linguistiques lexicaux⁴, que nous appellerons **mots-formes**. Dans le texte, le signifiant sera noté par l'italique et le signifié par les guillemets simples, appelés guillemets sémantiques. Ainsi, le mot-forme *cheval* véhicule le sens 'cheval'.

Une troisième dimension du signe linguistique demande à être présentée : il s'agit des **propriétés de combinatoire** – que la théorie Sens-Texte appelle le **syntactique** – qui

⁴ En effet, il existe d'autres types de signes linguistiques. Le suffixe du pluriel est aussi un signe linguistique dont le signifiant est (souvent) *-s* et dont le signifié ou sens est 'pluriel'. L'étude de tels signes linguistiques relève plutôt de la morphologie que de la lexicologie.

sont « les contraintes propres à ce signe qui limitent sa capacité à se combiner avec d'autres signes linguistiques et qui ne peuvent être déduites ni de son signifié, ni de son signifiant » (Polguère, 2003 : 37). Par exemple, la partie du discours⁵ fait partie des propriétés de combinatoire et fournit des indications quant aux combinaisons possibles avec d'autres signes linguistiques; un nom commun peut être précédé d'un article, il possède un genre et les adjectifs qui s'y rapportent doivent s'accorder en conséquence. Nous verrons plus tard que la connaissance du syntactique a une portée immense sur la maîtrise d'une unité lexicale, notamment en ce qui concerne la cooccurrence lexicale restreinte (voir II-1.5.3).

1.3.2 LE LEXÈME COMME REGROUPEMENT DE MOTS-FORMES

Nous considérerons donc *cheval* et *chevaux* comme deux mots-formes différents, mais nous serions tentés de dire qu'il s'agit en fait du même... mot !

Les deux mots-formes *cheval* et *chevaux* appartiennent à une entité abstraite que nous désignerons par le terme **lexème**. Le lexème, qui est un type particulier de **lexie** ou d'**unité lexicale**, est l'association signifié ~ signifiant stockée dans notre lexique mental qui nous permet de relier la suite de sons /ʃəval/ et le sens 'grand équidé à crinière'. Il faut mentionner que le concept de lexème dénote une entité abstraite et fait partie des termes du métalangage lexical ; seuls les mots-formes seront observables. Le nom d'une unité lexicale sera toujours mentionné en petites majuscules dans ce texte.

On peut donc dire que le lexème CHEVAL signifie 'cheval' et regroupe les mots-formes *cheval* et *chevaux*. De la même façon, le lexème ALLER signifie 'aller' et regroupe l'ensemble des formes fléchies de ce verbe (*vais, allons, iront, allas*, etc.). Plus généralement, le lexème est un regroupement de mots-formes qui ne se distinguent que par la flexion et dont le sens se retrouve dans le signifié de chacun de ces mots-formes.

⁵ La didactique du français semble privilégier les termes *classe de mots* ou *catégorie grammaticale* à celui de *partie du discours*. Pour des raisons qu'il serait trop long d'exposer ici, nous choisissons quand même le terme *partie du discours*. Pour une argumentation en faveur de ce choix, voir Polguère (2003, 78-80).

1.3.3 ET MONTER SUR SES GRANDS CHEVAUX?

Nous avons mentionné que le lexème est un type particulier d'unités lexicales. Observons maintenant un autre type. Certaines expressions dans la langue véhiculent un sens qui ne découle pas du sens de chacune des lexies les constituant. Ce phénomène s'appelle la **non-compositionalité sémantique**⁶ et nous désignerons par les termes **expressions idiomatiques** ou **locutions** de telles expressions.

Évidemment, de façon compositionnelle, nous pourrions comprendre l'expression *monter sur ses grands chevaux* en combinant les sens de chacune des lexies la constituant. L'expression signifierait ainsi 'se hisser sur le dos de chevaux de grande taille dont on est le propriétaire'. Bien sûr, aucun locuteur natif du français ne comprendra l'expression de cette façon, à moins de ne l'avoir jamais entendue; il comprendra plutôt 's'emporter, parler avec autorité et prétention'. Il n'est donc ici question ni de 'monter', ni de 'chevaux' et encore moins de la 'grande taille' de ces chevaux.

Comme pour les lexèmes, nous considérerons donc ces expressions linguistiques comme des unités lexicales de la langue. Ainsi, il existe en français une unité lexicale (de type locution) **MONTER SUR SES GRANDS CHEVAUX** qui signifie approximativement 'se mettre en colère' et regroupe les **expressions linguistiques**⁷ *monte sur ses grands chevaux*, *est monté sur ses grands chevaux*, etc. Les locutions sont omniprésentes dans les langues et peuvent appartenir à différentes parties du discours :

- locutions verbales : BRISER LA GLACE, NE PAS AVOIR FROID AUX YEUX, PRENDRE SON PIED, etc.
- locutions nominales : MAÎTRE CHANTEUR, CHAMBRE DES COMMUNES, etc.
- locutions adjectivales : MORT DE RIRE, BON MARCHE, etc.
- locutions adverbiales : EN RETARD, À L'ENVERS, À JAMAIS, etc.
- locutions prépositionnelles : À L'INSTAR (de), QUANT À, etc.

⁶ Une expression est **sémantiquement compositionnelle** lorsque son sens peut être déduit à partir du sens de chacune des lexies qui la composent. Par exemple, la phrase *Il achète une pomme* signifie 'il'+ 'achète'+ 'pomme'. La **non-compositionalité sémantique** se manifeste quant à elle dans les expressions idiomatiques pour lesquelles les sens de chacune des lexies les composant ne permettent pas de saisir leur sens. Par exemple, la locution *BAYER AUX CORNEILLES* signifie 'perdre son temps en regardant en l'air niaisement' et son sens ne contient aucun des sens des lexies qui la composent. Pour une présentation plus détaillée, voir Mel'čuk, 2003 : 20-21.

⁷ Par **expression linguistique**, nous entendons tout groupe de mots-formes qui forme un syntagme. Par exemple, dans la phrase *Le chien brun du voisin a saccagé mes platebandes*, le groupe *le chien brun du voisin* ou *mes platebandes* constituent des expressions linguistiques, alors que les groupes *brun du voisin* ou *a saccagé mes* ne seront pas considérés comme tels, puisque, ne correspondant pas à des « unités » participant à la structure syntaxique de la phrase, ils ne forment pas des syntagmes.

1.3.4 DÉFINITION DE LA NOTION D'UNITÉ LEXICALE (LEXIE)

Comme nous avons vu qu'il existe deux types de lexies de nature différente, deux termes différents seront utilisés pour les désigner. Voici donc la définition de lexie, telle que proposée par Polguère (2003 : 50) :

Une **lexie**, aussi appelée **unité lexicale**, est un regroupement 1) de mots-formes ou 2) de constructions linguistiques que seule distingue la flexion. Dans le premier cas, il s'agit de **lexèmes**, dans le second cas, de **locutions**.

Chaque lexie (lexème ou locution) est associée à un sens donné, que l'on retrouve dans le signifié de chacun des signes (mots-formes ou constructions linguistiques) auxquels elle correspond.

1.3.5 LES LEXIES AYANT UN SIGNIFIANT IDENTIQUE

Comme les sens à exprimer sont infinis, mais que les langues possèdent un nombre fini de phonèmes et ne sauraient s'encombrer de signifiants extrêmement longs et difficiles à mémoriser, il sera fréquent d'avoir un même signifiant pour plusieurs signifiés. En fait, la majorité des « mots » courants d'une langue ont plus d'un sens. Comme nous savons que chaque lexie est associée à un sens donné et un seul, nous pouvons donc déduire qu'il y aura plusieurs lexies partageant un signifiant identique. Deux cas de figure sont possibles.

Tout d'abord, plusieurs lexies partagent un même signifiant et ont des sens qui présentent des liens évidents. Par exemple, les lexies BRÛLER₁ [*La feuille brûle.*] et BRÛLER₂ [*Il brûle la feuille.*] partagent clairement un lien de sens puisque BRÛLER₂ est le causatif de BRÛLER₁ (*Il brûle₂ la feuille* = 'il cause que la feuille brûle₁'). De la même façon, le lien de métaphore est incontestable entre les lexies VOLER¹₁ [*Les oiseaux volent dans le ciel.*] (= 'se déplacer dans les airs') et VOLER¹₂ [*La ballerine volait sur la scène.*] (= 'se déplacer rapidement, avec légèreté'). Nous parlerons alors de **polysémie** et nous regrouperons ces lexies sous une même entité que nous appellerons **vocab**. Chacune des lexies constituant le vocable polysémique⁸ sera désignée par le terme d'**acception**. Considérant ces deux seuls sens de VOLER, nous dirons qu'il existe un vocable polysémique VOLER à deux acceptions : VOLER¹₁ et VOLER¹₂.

⁸ Il existe bien sûr des vocables monosémiques, c'est-à-dire dont la forme n'est associée qu'à une seule unité lexicale; il s'agit d'un vocable à une seule acception. C'est le cas du vocable ORNITHORYNQUE, par exemple.

Le deuxième cas de figure serait celui où deux ou plusieurs lexies partagent un même signifiant sans qu'il n'existe pourtant de lien de sens entre elles. Par exemple, il ne semble pas y avoir de lien de sens entre VOLER² [*Le voyou vole le sac de la dame.*] et VOLER¹₁ [*Les oiseaux volent dans le ciel.*]. Il s'agit alors d'un cas d'**homonymie**.

La polysémie est une propriété des vocables, alors que l'homonymie est une relation entre vocables; par conséquent, des vocables homonymes pourront tout à fait être polysémiques (c'est d'ailleurs le cas de VOLER¹, qui regroupe les lexies VOLER¹₁ et VOLER¹₂).

1.3.6 LA SYNONYMIE ET LA PARAPHRASE

Nous avons vu des cas où les lexies avaient des signifiés différents (apparentés ou non), mais partageaient des signifiants identiques⁹. Nous pouvons donc imaginer l'inverse : des lexies ayant des signifiants différents, mais dont les signifiés seraient identiques, ou du moins similaires. Il s'agit du cas bien connu de **synonymie**. La **synonymie exacte** est plutôt rare dans les langues puisqu'il est somme toute assez peu utile d'avoir deux lexies différentes pour désigner une même réalité. Il existe quand même quelques cas de synonymie exacte : AUTOMOBILE ~ VOITURE, VÉLO ~ BICYCLETTE, AN ~ ANNÉE, JOB ~ BOULOT, etc. La **synonymie approximative** est beaucoup plus répandue. Polguère (2003 : 122) la définit comme la relation de deux lexies entre lesquelles « il y a soit intersection soit inclusion de sens telle que les deux lexies L₁ et L₂ peuvent être considérées comme dotées d'une valeur sémantique suffisamment proche pour que l'on puisse, en utilisant l'une ou l'autre, exprimer sensiblement la même chose ». La synonymie approximative est celle que l'on retrouve dans la plupart des dictionnaires commerciaux. Ce phénomène existe aussi au niveau des expressions linguistiques : « deux expressions linguistiques ayant (approximativement) le même sens sont appelées des **paraphrases** » (Polguère, 2003 : 98).

Il nous paraît intéressant d'introduire la notion de synonymie pour mettre en lumière le fait que, même si les sens de deux lexies synonymes sont très proches, chacune aura ses propres propriétés de combinatoire et qu'on ne peut remplacer librement une unité lexicale par un synonyme dans tous les contextes. Par exemple, les lexèmes TRISTE et

⁹ Nous parlons de signifiants identiques au pluriel, car il faut se rappeler qu'à une lexie peuvent correspondre plusieurs signifiants. Par exemple, les lexies BRÛLER₁ [*La feuille brûle.*] et BRÛLER₂ [*Il brûle la feuille.*] ont des signifiés différents, mais partagent les mêmes signifiants *brûler*, *brûle*, *brûleras*, *brûleront*, etc.

MALHEUREUX sont des synonymes, mais leurs collocations sont différentes. Par exemple, pour intensifier TRISTE OU MALHEUREUX, nous emploierons les expressions *triste à mourir* et *malheureux comme les pierres*, mais pas **triste comme les pierres* et **malheureux à mourir*.

1.4 LE SENS ET LA COMBINATOIRE SÉMANTIQUE

La propriété centrale des unités lexicales est leur signifié, puisque celles-ci existent précisément pour véhiculer un sens. Pourtant, même si nous utilisons sans trop y réfléchir des unités lexicales pour exprimer des messages, il est très difficile de définir clairement ce qu'est le sens. Quels sont les éléments que nous retrouvons dans le sens d'une unité lexicale?

La définition lexicographique est la manière typique de décrire le sens des unités lexicales. Elle consiste en fait à en donner une paraphrase, c'est-à-dire à fournir une expression qui véhicule le même sens que la lexie décrite. Cette paraphrase, que la théorie Sens-Texte appelle décomposition sémantique¹⁰, est constituée de termes sémantiquement plus simples que celui décrit et se divise en deux parties : le genre prochain et les différences spécifiques.

Le genre prochain est la composante de sens centrale de la définition, et constitue « en quelque sorte une paraphrase approximative minimale de la lexie définie. » (Polguère, 2003 : 150) Par exemple, le genre prochain de 'canard' serait 'oiseau'¹¹.

La seconde partie de la définition regroupe les différences spécifiques, qui sont des informations qui caractérisent le genre prochain de sorte à ce qu'il corresponde au sens de l'unité lexicale décrite en le différenciant des autres unités lexicales qui partageraient le même genre prochain. Pour 'canard', les différences spécifiques seraient les suivantes : 'palmipède', 'aquatique', 'au bec large', 'au cri nasillard', 'dont certaines espèces sont domestiques' (Antidote Prisme, 2004 : article CANARD) ; elles permettent à l'unité lexicale CANARD de ne dénoter que l'oiseau de cette espèce et aucun autre.

¹⁰ La décomposition sémantique consiste à exprimer un sens par des sens plus simples. Mel'čuk (1997) parle de « formules moléculaires » du sens. Cette perspective suppose donc que la plupart des sens d'une langue sont décomposables, à part un nombre limité de sens, qui constituent des primitifs sémantiques. Pour plus de détail, voir Mel'čuk *et al.* (1995).

¹¹ Souvent, le genre prochain correspond à un hyperonyme de la lexie décrite.

En plus du genre prochain et des différences spécifiques, qui constituent une méthode de définition assez traditionnelle, les décompositions sémantiques de la théorie Sens-Texte spécifient aussi le typage des actants. Ce typage des actants consiste à mettre des contraintes sémantiques sur les participants de la situation dénotée par l'unité lexicale décrite. Par exemple, le verbe MONTER [*L'enfant monte à un arbre.*] qui, dans cette acception, signifie 'se déplacer dans un mouvement de bas en haut' nécessite que son sujet dénote un être animé. En rendant compte de cette particularité dans la définition, nous limitons les unités lexicales qui pourront apparaître en cooccurrence avec ce verbe.

Tous ces éléments du sens auront évidemment une influence sur l'agencement des lexies entre elles; c'est que nous appelons la combinatoire sémantique. Chaque unité lexicale impose en fait des conditions d'utilisation découlant de son sens. Par exemple, le verbe PARLER exige que le nom qui jouera auprès de lui le rôle de sujet dénote un être vivant qui puisse s'exprimer par la parole. Ainsi, tous les noms qui dénotent des objets, des animaux ou des concepts ne pourront servir de sujet à PARLER. De la même façon, la lexie CASSER demande que son complément d'objet direct dénote un objet rigide et il sera étrange de dire **L'homme a cassé sa chemise en tombant*. Par contre, la phrase *L'homme a cassé son verre en tombant* est tout à fait correcte, puisque les composantes de sens 'objet' et 'rigide' de VERRE satisfont le typage du deuxième actant sémantique de CASSER.

Donc, par la décomposition sémantique et le typage des actants, nous pouvons rendre compte des conditions sémantiques qui restreignent en quelque sorte la **cooccurrence lexicale**¹². Nous verrons à la section suivante d'autres types de combinaisons qui, cette fois, ne peuvent pas être déduites à partir du sens d'une unité lexicale.

1.5 LES PROPRIÉTÉS DE COMBINATOIRE

En plus du signifiant et du signifié, nous considérons, comme nous l'avons mentionné plus haut, qu'il existe une troisième dimension au signe linguistique en général, et à la lexie en particulier : les propriétés de combinatoire. Ces propriétés de combinatoire ont été définies comme « les contraintes propres à ce signe qui limitent sa capacité à se combiner avec d'autres signes linguistiques et qui ne peuvent être déduites ni de son

¹² La *cooccurrence lexicale* est le simple fait pour deux ou plusieurs lexies d'apparaître ensemble dans un même énoncé.

signifié, ni de son signifiant » (Polguère, 2003 : 37). Nous distinguerons deux axes de combinatoire¹³, bien qu'ils soient intimement reliés : la **combinatoire grammaticale** – qui regroupe la combinatoire syntaxique ou **régime** ainsi que d'autres propriétés inhérentes à l'unité lexicale – et la **cooccurrence lexicale restreinte**.

1.5.1 LA COMBINATOIRE SYNTAXIQUE OU RÉGIME

Le sens de beaucoup de lexies du français implique des participants. Par exemple, la lexie **INSULTER** signifie 'qqn. insulte qqn.'. Nous appelons **lexies prédicatives** ou **prédicats sémantiques** de telles unités lexicales, et **arguments** ou **actants sémantiques** les participants impliqués dans la situation qui s'expriment dans des structures particulières auprès de leur prédicat¹⁴. Ces structures sont appelées **structures prédicatives** ou **actanciell**es. Tous les verbes sont des prédicats, mais aussi tous les adjectifs, les adverbes et certains noms. En effet, **GRAND** signifie 'qqn./qqch. est grand', **LENTEMENT** signifie 'qqch. se déroule lentement' et **FRÈRE** signifie 'qqn. est le frère de qqn.'. Le nombre d'arguments varie bien sûr d'un prédicat à l'autre, mais le nombre maximal semble être de six.

Pour les verbes, ce phénomène se traduit dans la grammaire par les notions de **transitivité** et d'**intransitivité**. Ainsi, un verbe ayant seulement un sujet comme argument sera appelé **verbe intransitif** (**BÂILLER** = 'qqn. bâille'). Un verbe ayant plus d'un argument sera appelé **verbe transitif** (**CONNAÎTRE** = 'qqn. connaît qqn./qqch.', **ATTRIBUER** = 'qqn. attribue qqch. à qqn.'). Certains verbes transitifs pourront tolérer que leur complément d'objet ne soit pas exprimé ; par exemple, on peut dire *L'homme mange* même si **MANGER** = 'qqn. mange qqch.'. Nous parlerons alors d'**emploi absolu** du verbe transitif.

Beaucoup de verbes, ainsi que des noms et certains adjectifs, introduisent leurs arguments par des prépositions qui sont vides de sens, et qui n'ont pour fonction que

¹³ C'est tout à fait délibérément que nous ne considérons pas la combinatoire sémantique parmi les propriétés de combinatoire. En effet, les propriétés de combinatoire de l'unité lexicale « ne peuvent être déduites ni de son signifié, ni de son signifiant » (Polguère, 2003 : 37). Or, nous venons de voir à la section II-1.4 que la combinatoire sémantique est tout à fait prévisible à partir de la définition (donc du signifié) d'une lexie.

¹⁴ La notion de prédicat sémantique permet de diviser les sens lexicaux en deux grandes classes : d'une part, les **prédicats sémantiques** dont le sens inclut des participants (ou arguments), comme 'émettre', 'souci' ou 'rapidement' ; d'autre part, les **objets sémantiques** dont le sens n'implique pas de participants, comme 'nuage', 'radis', 'lundi', etc. (Polguère, 2003 : 108)

de relier ces lexies à leurs compléments; ce sont les **prépositions régies**. Par exemple, *qqn. parle à qqn. de qqch., acquiescement de qqn. à qqch., fier de qqch., etc.*

La transitivité ou l'intransitivité, le caractère obligatoire ou non des compléments, le recours aux prépositions régies, voilà autant de propriétés inhérentes à chaque unité lexicale. Ces propriétés relèvent bien sûr de la syntaxe, mais non pas de la syntaxe grammaticale de la langue. Il s'agit d'une syntaxe à un autre niveau : une syntaxe lexicale ou microsyntaxe. On a tendance, à tort, à considérer ces phénomènes comme des questions de syntaxe; cette microsyntaxe étant soumise aux « caprices » de chaque lexie, elle doit par conséquent être considérée comme une des facettes du lexique.

1.5.2 LES AUTRES PROPRIÉTÉS DE LA COMBINATOIRE GRAMMATICALE

D'autres propriétés inhérentes à l'unité lexicale régiront son comportement grammatical. Le genre du nom est un exemple de ces propriétés propres à la lexie ; il ne peut être déduit ni de son signifiant, ni de son signifié, et a un impact sur sa façon de se combiner avec d'autres signes linguistiques, puisque c'est cette caractéristique qui déterminera comment accorder les adjectifs et les déterminants accompagnant ce nom. La partie du discours est une autre propriété de combinatoire grammaticale qui nous informera, notamment, du rôle syntaxique que peut jouer l'unité lexicale.

Il existe aussi des propriétés de combinatoire grammaticale moins fréquentes. Par exemple, certaines unités lexicales ont la particularité de n'exister qu'au singulier (CÉLIBAT [*Le célibat des prêtres*], AGNEAU [*L'agneau est une viande très tendre.*]) ou au pluriel (FUNÉRAILLES, PRÉMICES, VACANCES). Certains verbes, appelés défectifs, ne se conjuguent pas à certains temps (FRIRE, PAÎTRE). Des lexies spécifiques imposent l'emploi du subjonctif dans la subordonnée qu'elles introduisent (SOUHAITER, AVANT QUE). Certaines locutions ne tolèrent pas la mise au passif (FORCER LA MAIN; *Ses supérieurs lui ont forcé la main*, mais pas **La main lui a été forcée par ses supérieurs*). Voilà donc autant d'exemples de propriétés de combinatoire grammaticale dont il faut tenir compte dans l'emploi d'une unité lexicale.

1.5.3 LA COOCCURRENCE LEXICALE RESTREINTE : LES COLLOCATIONS

Nous avons vu que certaines conditions s'appliquent dans le cas de la cooccurrence lexicale. Pourtant, certaines lexies semblent entretenir des liens privilégiés. En effet, il existe un autre axe de propriétés de combinatoire : la **cooccurrence lexicale restreinte**. Cette cooccurrence lexicale restreinte se manifeste dans la langue par les **collocations** ou **expressions semi-idiomatiques**. Le terme semi-idiomatique renvoie au principe de compositionnalité sémantique mentionné plus haut (II-1.3.3). Polguère (2003 : 134) définit la collocation comme suit :

L'expression *AB* (ou *BA*), formée des lexies *A* et *B*, est une collocation si, pour produire cette expression, le locuteur sélectionne *A* librement d'après son sens '*A*', alors qu'il sélectionne *B* pour exprimer un sens '*C*' en fonction de *A*.

Nous parlons d'expressions semi-idiomatiques parce qu'une lexie, que l'on choisit pour exprimer un certain sens, nous impose des choix lexicaux pour exprimer auprès d'elle d'autres sens. Par exemple, pour exprimer 'quelqu'un dort', nous utiliserons le verbe DORMIR. Mais, pour signifier que 'son sommeil est intense', le verbe DORMIR nous contraindra à utiliser certaines lexies et nous produirons les collocations *dormir profondément, comme une bûche, comme un bébé, à poings fermés*, etc. Par contre, si nous choisissons le verbe RIRE plutôt que DORMIR et que nous voulons exprimer encore une fois le sens de 'rire intensément', nous aurons accès aux collocations *rire aux larmes, à gorge déployée, aux éclats*, etc.

Nous dirons que la lexie choisie librement est la **base de la collocation**, alors que les expressions contraintes seront appelées **collocatifs**. Évidemment, chaque lexie appelle des collocatifs différents et il serait aberrant de dire que **Le bébé dormait aux éclats* ou que **Le jeune homme riait comme une bûche*, même si les deux collocatifs véhiculent ici le même sens d'intensification.

Il existe différents types de collocations et il est impossible d'en faire ici une présentation exhaustive. La théorie Sens-Texte, à l'aide des fonctions lexicales, propose une méthode d'analyse et de modélisation formelles des collocations (voir Mel'čuk *et al.*, 1995).

Leur nombre étant extrêmement élevé, les collocations sont omniprésentes dans la langue et sont généralement l'apanage du locuteur natif; une manipulation maladroite des collocations trahit souvent le locuteur d'une langue seconde.

CONCLUSION

La présente section a présenté les concepts de base de l'étude du lexique. Nous comprenons maintenant que la connaissance d'une unité lexicale est complexe et multidimensionnelle. Deux choses importantes ressortent de cette section :

- le terme *mot* est ambigu et ne permet pas un discours précis sur le lexique ; il doit être remplacé par les termes *mot-forme*, *lexie* (= *unité lexicale*) ou *vocable* selon le cas;
- il existe différents éléments constitutifs du sens d'une unité lexicale;
- les propriétés de combinatoire occupent une place immense dans la maîtrise d'une unité lexicale.

Armé de ces notions et de cette terminologie propres à la lexicologie, nous pouvons maintenant entreprendre notre définition de l'erreur lexicale.

2. CARACTÉRISATION DE LA NOTION D'ERREUR LEXICALE

Le premier objectif de notre recherche vise à décrire les types d'erreurs lexicales commises par les étudiants. Si nous avons décrit les principaux concepts linguistiques auxquels nous ferons appel dans nos analyses lexicales, il importe maintenant de définir clairement ce que nous entendons par « erreur lexicale ». Nous avons observé dans notre problématique (voir I-3.3) qu'un certain flou plane au-dessus de cette notion d'erreur lexicale. Nous discuterons dans la section suivante de la notion de norme, dont on ne peut faire l'économie lorsqu'on aborde le thème de l'erreur, puis nous déterminerons de quelle façon la présente recherche traitera la norme. Nous verrons ensuite de quelle façon cette norme s'applique au lexique en reprenant les notions présentées dans la section précédente. Nous tenterons aussi de démontrer pourquoi il est nécessaire, selon nous, d'élargir le concept d'erreur lexicale pour qu'il englobe certains dysfonctionnements qui pourraient à première vue être considérés comme relevant d'autres dimensions du code linguistique. Cette exploration de la façon dont la norme s'applique au lexique nous permettra de développer, dans notre méthodologie (voir III-2.5), un outil servant à caractériser précisément les différents problèmes de nature lexicale.

2.1 L'ERREUR ET LA NORME

Comme nous l'avons vu plus haut, la langue peut être fonctionnellement décrite en tant que système établissant une correspondance entre des sens et des textes. Ce système permet d'abord et avant tout aux membres d'une même communauté linguistique de communiquer entre eux, oralement ou par écrit. Évidemment, pour que la communication soit possible, la langue doit être *stable* et *codifiée*¹⁵. En effet, c'est précisément parce qu'ils partagent un code commun que deux humains peuvent communiquer. « Il existe dans toute langue des *invariants* qui sont inhérents à la *structure* de la langue elle-même, qui sous-tendent les *systèmes* de l'oral et de l'écrit qui lui assurent la stabilité [...], et ceci dans un but essentiel : l'efficacité de la

¹⁵ Saussure parle de l'immutabilité du signe linguistique, qui assure au code linguistique une stabilité qui permet d'utiliser une langue ou d'en apprendre de nouvelles. Paradoxalement, une autre des caractéristiques proposées par Saussure est la mutabilité du signe linguistique, qui permet à la langue d'évoluer. La mutabilité du signe linguistique se manifeste notamment par deux types de variations : les variations individuelles et les transformations dans le temps (Polguère, 2003 : 35).

communication » (Honveau, 1995 : 53). Corbeil (1980 : 47) et Chaudenson, Mougeon et Beniak (1993 : 7) désignent cette zone d'invariance de la langue par le nom de « noyau dur ».

Au-delà de ces invariants, nous savons que l'usage de la langue, même au sein d'une communauté linguistique, présente de nombreuses variations. Les sources de variations peuvent être de natures diverses : diachronie, mode de communication, situation géographique, appartenance sociale, domaine technique, etc. (Polguère, 2003 : 74). Les variations peuvent s'observer aux niveaux phonique et lexical, et, dans une moindre mesure, aux niveaux syntaxique et morphologique. Ces variations sont évidemment beaucoup plus présentes à l'oral qu'à l'écrit, le code écrit étant plus rigide et assujéti à la norme. « La variation n'est absolument pas un obstacle à la communication tant que les relations nécessaires au fonctionnement du système [les invariants] sont préservées » (Honveau, 1995 : 55).

Le concept de *langue* est étroitement lié à celui de *norme*. Bien entendu, il existe différentes conceptions de ce qu'est ou doit être la norme. Nous en présenterons ici trois conceptions différentes; ce sont à ces normes que nous ferons appel pour catégoriser un problème lexical.

Le *Dictionnaire de linguistique et des sciences du langage* (Dubois, 1994) caractérise la norme comme un « système d'instructions définissant ce qui doit être choisi parmi les usages d'une langue donnée si l'on veut se conformer à un certain idéal esthétique ou socio-culturel. » Nous pourrions qualifier cette approche de **norme prescriptive** ; elle édicte les usages qui doivent être conservés parmi les variantes linguistiquement possibles et condamne par le fait même les autres. Bien qu'elle évoque certains principes tels la racine historique, l'accord avec les habitudes de la langue ou la justification « logique » (Ducrot et Schaeffer, 1995 : 312), la norme relève souvent d'un choix arbitraire soumis à des enjeux « politiques » ou à la fantaisie de certains grammairiens – à ce sujet, le livre *Les fautes de français existent-elles ?* de Danielle Leeman-Bouix (1994) est éloquent. Paradoxalement, d'un point de vue diachronique, on s'aperçoit que c'est souvent le parler familier ou argotique – donc hors norme d'un point de vue prescriptif – d'une époque qui déterminera la norme de l'époque suivante. La norme prescriptive, qui alimente la grammaire normative d'une langue et en balise l'usage, est d'autant plus présente en français qu'il existe des organismes comme l'Académie française en France ou l'Office québécois de la langue française au Québec qui ont autorité en ce domaine. Selon cette norme, une *erreur* est commise

chaque fois qu'il y a écart par rapport à la variante privilégiée parmi celles possibles, mêmes celles qui seraient linguistiquement justifiables. Cette norme prescriptive frappe particulièrement le code écrit, surtout au niveau des accords grammaticaux et de l'orthographe d'usage ; ainsi, il sera défendu d'écrire **nèje* pour *neige*, *jeu* n'admettra que *x* et non *s* pour marquer son pluriel et le verbe *dire* à la 2^e personne du pluriel du présent de l'indicatif ne tolérera pas d'être actualisé par **vous disez* (ce qui serait par ailleurs davantage en accord avec le système de la langue). La norme prescriptive ne fait pas vraiment de distinction entre l'expression écrite et l'expression orale ; la langue écrite étant la « bonne », l'oral doit la refléter.

Les linguistes conçoivent souvent la norme d'une façon fort différente. Pour eux, c'est l'usage qui la détermine ; c'est donc par l'analyse de la parole au sens saussurien que les linguistes dégageront les formes courantes, donc « normales » de l'usage. On pourrait qualifier cette approche de **norme descriptive**. La norme est ainsi fixée par les locuteurs et change en fonction des variations géographiques, sociales, etc. Par conséquent, la norme écrite différera de la norme orale. La norme descriptive étant statistique, une erreur sera donc un écart par rapport à l'usage de la majorité. Ainsi, *une autobus* ou *un aréoport* pourraient être normés dans certains milieux.

Un autre type de norme recoupe les deux autres de façon transversale : il s'agit de la **norme fonctionnelle**. Selon ce point de vue, c'est le contexte qui « impose » la norme ; la norme écrite sera donc différente de la norme parlée, qui elle-même changera selon que la conversation a lieu dans un café, lors d'une réunion avec un supérieur, dans les vestiaires, etc. « Dans la perspective fonctionnelle, le bon usage est celui qui est adapté aux diverses situations de communication, aux situations de la vie en société » (Dallaire, 1997 : 30). Ainsi, un parlé « proustien » lors d'une partie de pêche sera aussi déviant par rapport à la norme que le tutoiement lors d'une rencontre avec le premier ministre. Il y a donc « erreur » si la variété ou le registre de langue n'est pas en adéquation avec le contexte de l'acte langagier. Évidemment, la norme fonctionnelle se marie peut-être mieux à la norme descriptive qu'à la norme prescriptive « pure », qui voudrait plutôt voir l'utilisation, en toutes circonstances, d'un français uniforme calqué sur le code écrit.

2.2 LA PRÉSENTE RECHERCHE ET LA NORME

La présente recherche s'intéresse à l'erreur lexicale, et demande donc de s'intéresser à la notion de norme. En effet, le concept même d'erreur implique nécessairement une idée de déviation, qui sera elle-même observée par rapport à une norme. Une fois qu'il est admis qu'il n'existe pas qu'une norme et que cette notion a plutôt tendance à se conjuguer au pluriel, il est maintenant important de déterminer où se situera la norme dans le cadre de notre recherche. La perspective fonctionnelle nous permettra d'abord d'opérer certains choix.

Notre corpus. Les erreurs lexicales seront recensées dans un corpus de textes rédigés dans le cadre d'un cours universitaire. Conséquemment, le français employé dans ces textes devrait être un français écrit courant ou soutenu; les registres de langue familier ou vulgaire seront donc considérés comme déviants par rapport au contexte.

La présente recherche se situe dans une perspective didactique. Comme ce mémoire est élaboré dans le cadre d'une maîtrise en didactique du français et que nous aimerions que notre typologie d'erreurs lexicales puisse être réutilisée par des enseignants en classe, il convient d'adopter ici une approche « scolaire ». Nous savons que le français écrit est très rigide en ce qui concerne l'orthographe grammaticale. En ce domaine, comme en syntaxe, la grammaire impose ses règles de façon systématique et tout scripteur doit s'y conformer; c'est la norme prescriptive qui domine. Pour le lexique, par contre, nous verrons que, en fonction de l'aspect de l'unité lexicale qui nous intéresse, la norme à considérer ne sera pas toujours la même, puisque la norme touchant certaines dimensions du lexique est définie par l'usage plutôt que par des règles strictes. Malgré tout, certains aspects du lexique pourront être analysés dans une perspective prescriptive, et nous adopterons cette approche pour deux raisons; d'abord, pour être en accord avec le monde scolaire et, ensuite, pour tenter de minimiser une part de la subjectivité inhérente au concept même d'erreur lexicale en utilisant des ouvrages de référence qui trancheront à propos du statut erroné ou non du problème lexical dans les cas litigieux.

Quelques remarques méritent d'ailleurs d'être faites à propos de notre concept d'erreur. Loin de nous l'idée de condamner purement et simplement un usage considéré fautif; les erreurs relevées dans cette recherche constituent des erreurs dans le présent contexte, mais pourraient être tout à fait acceptables dans d'autres conditions.

Par contre, cela ne remet pas en doute la qualité ou la pertinence de notre typologie des problèmes lexicaux, puisque celle-ci pourra être réutilisée dans d'autres contextes en déterminant une nouvelle norme. Il est important pour nous de pouvoir trancher à propos du statut erroné ou non d'un problème lexical (voir III-2.5.3) afin de poursuivre cette recherche; la même situation existe d'ailleurs pour un enseignant en situation de correction de textes. Par contre, au-delà de la décision de considérer comme une erreur ou non un tel problème, une autre question nous semble plus importante : si nous admettons ce problème comme erreur, comment notre typologie peut-elle en rendre compte?

2.3 LE LEXIQUE ET LA NORME

Par l'exploration des différentes dimensions de l'unité lexicale présentée à la section II-1, nous tenterons de voir de quelle façon la norme peut s'y appliquer, et selon quels cas de figure cette dernière pourra être transgressée. Cette démarche nous permettra de définir notre conception de l'erreur lexicale. La section II-2.3.1 justifiera notre conception « large » de la notion d'erreur lexicale. La section II-2.3.2 s'intéressera aux dimensions de la lexie pour lesquelles la norme est assez claire et consensuelle, alors que le point II-2.3.3 sera plutôt consacré aux aspects du lexique qui semblent plus difficiles à caractériser en fonction d'une norme prescriptive. Finalement, la section II-2.3.4 présentera un type particulier de problème lexical : l'utilisation d'une lexie fictive.

2.3.1 LES DIFFÉRENTS VISAGES DE L'ERREUR LEXICALE

À la lecture des points II-2.3.2 et II-2.3.3, certains pourront être choqués de constater que notre conception de l'erreur lexicale recouvre plusieurs dysfonctionnements qu'ils auraient plutôt considérés comme relevant d'autres dimensions du code. Ainsi, peut-être serons-nous accusé de « ratisser un peu large » en matière d'erreur lexicale. C'est pourquoi nous croyons qu'une mise au point s'impose avant même d'aborder cette question.

Nous considérons comme lexicale toute information relative à une lexie en particulier. Tous s'entendront pour dire, par exemple, que l'orthographe ou le sens d'une unité

lexicale sont des propriétés qui lui sont inhérentes. Mais nous avons vu, dans notre section sur le lexique (voir II-1.5), que les propriétés de combinatoire d'une unité lexicale lui sont aussi particulières. Cette idiosyncrasie devrait donc nous convaincre qu'un dysfonctionnement découlant de la transgression d'une propriété de combinatoire devra être considéré comme lexical. En effet, les règles de grammaire de la langue ne seront pas en mesure d'en rendre compte sans faire appel à une particularité d'une unité lexicale spécifique.

Prenons par exemple le cas du genre du nom. Cette caractéristique est complètement arbitraire (nous ne pouvons la déduire d'aucune façon) et propre à chaque nom ; c'est d'ailleurs la raison pour laquelle elle figure dans le dictionnaire. Par conséquent, il faut donc admettre que le genre du nom est une propriété lexicale. Maintenant, si nous sommes confrontés au syntagme suivant : **faire un escapade*, il faudra considérer que nous sommes devant une erreur lexicale puisque le nom féminin ESCAPADE est ici utilisé comme un nom masculin, ce qui viole une de ses propriétés inhérentes. De la même façon, il faudra analyser **un encyclopédie volumineux* comme une erreur lexicale puisque, encore une fois, c'est du genre du nom, une propriété lexicale, que découle l'erreur. Certains prétendront plutôt qu'il s'agit d'une erreur grammaticale, car l'accord avec l'adjectif VOLUMINEUX est fautif. Mais comme le déterminant aussi est au masculin, n'y a-t-il pas plutôt lieu de croire que l'étudiant qui écrit **un encyclopédie volumineux* a appliqué de façon impeccable les accords grammaticaux, et que c'est en fait la propriété lexicale « genre féminin » qu'il ne connaissait pas ? Le fait qu'il soit extrêmement embarrassant d'affirmer que ce segment présente en fait deux erreurs d'accord – une avec le déterminant et une avec l'adjectif – milite aussi en faveur de notre position. Bien entendu, nous sommes tout à fait d'accord que ce dysfonctionnement d'origine lexicale se manifeste au niveau des accords grammaticaux. C'est pourquoi nous parlerons d'**erreur lexicogrammaticale**.

Nous pouvons démontrer la même chose du côté de la syntaxe. Par exemple, le verbe ENJOINDRE demande la construction *X enjoint à Y de faire Z*. Cette caractéristique du verbe est idiosyncrasique et aucune règle de syntaxe du français ne nous permettra de la déduire. Les prépositions servant à introduire les différents compléments du verbe font donc aussi partie des propriétés de combinatoire de l'unité lexicale. Ainsi, nous devrions considérer que la phrase **Il a enjoint son frère à participer au concours* présente des erreurs lexicales. Encore une fois, certains pourront affirmer qu'il s'agit plutôt de syntaxe, car il « manque un mot » devant *son frère*. Il s'agit en effet d'un

problème de syntaxe, mais de microsyntaxe ou de syntaxe lexicale (voir II-1.5.1). Nous croyons encore une fois qu'il faut éviter de parler ici d'erreur syntaxique, mais plutôt d'erreur **lexicosyntaxique**; la source de l'erreur est lexicale alors que sa manifestation est syntaxique.

Nous rappelons donc, avant de commencer notre exploration de la notion d'erreur lexicale, que nous considérerons comme lexicale toute erreur découlant de la violation d'une propriété inhérente à une lexie particulière. Bien entendu, cette violation aura des incidences sur d'autres dimensions du code, puisqu'elle se manifestera dans la parole, qui conjugue à la fois les dimensions lexicale, syntaxique et morphologique. De la même façon que le médecin diagnostique une maladie en observant les symptômes, nous pourrions remonter à la source lexicale d'un dysfonctionnement en apparence grammatical ou syntaxique pour comprendre plus justement la nature du problème. En effet, nul ne se confierait aux soins d'un médecin se contentant de constater d'inquiétantes plaques rouges sur le corps d'un patient sans chercher à en comprendre l'origine!

2.3.2 LE LEXIQUE ET LA NORME PRESCRIPTIVE

Les paragraphes suivants présenteront les aspects du lexique pour lesquels l'usage standard semble plutôt fixé. Un doute concernant ces différents aspects de l'unité lexicale pourra être levé en vérifiant dans les ouvrages de référence; il sera possible de statuer de façon assez catégorique sur un problème lexical relevant de ces aspects. C'est ici la norme prescriptive qui fera autorité.

Le signifiant :

Le **signifiant** de l'unité lexicale est certainement sa dimension la plus tangible, notamment à l'écrit, où il se manifeste par une suite de caractères. Cet aspect de la lexie est aussi très codifié et soumis aux règles strictes de l'orthographe lexicale, plus communément appelée *orthographe d'usage*.

L'**orthographe d'usage** est aussi figée que les règles grammaticales et ne laisse que peu de latitude au scripteur. Bien sûr, il existe de nombreux cas – bien qu'ils ne représentent qu'une infime partie du lexique français – où deux orthographes concurrentes sont acceptées, par exemple tous ceux touchés par les nouvelles rectifications de l'orthographe française (Réseau pour la nouvelle orthographe du

français, 2004). Par contre, ces cas particuliers ne témoignent pas selon nous d'une grande souplesse dans l'orthographe du français, mais découlent plutôt d'une période de transition due à une évolution normale de la langue. Si deux choix coexistent aujourd'hui, il est plus que probable qu'un seul subsistera et deviendra normé d'ici quelques années. C'est ici la norme prescriptive qui fait loi et l'orthographe lexicale acceptée sera celle qui figure dans les ouvrages de référence.

La norme étant extrêmement rigide en ce qui concerne le signifiant, il est facile de comprendre de quelle façon elle pourra être transgressée : en ne respectant pas l'orthographe lexicale recommandée. Nous distinguons deux cas.

Le premier est celui de l'**erreur d'orthographe d'usage**, qui indique une méconnaissance de la forme écrite du signifiant de la lexie employée. Par exemple, **chateau* et **collonne* présentent ce type d'erreur, car leur orthographe dévie de celle imposée par les ouvrages de référence.

Le deuxième cas de transgression du signifiant est le **barbarisme**. Ce type d'erreur trahit une méconnaissance du signifiant plus profonde qui se manifeste non seulement à l'écrit, mais aussi à l'oral. Autrement dit, c'est un signifiant déformé¹⁶ qui est associé au signifié de la lexie dans la tête de la personne commettant ce type d'erreur. Par exemple, **infractus* pour INFARCTUS ou **en bon uniforme* pour EN BONNE ET DUE FORME.

Le signifié :

Le **sens des unités lexicales** constitue évidemment leur propriété centrale, puisqu'elles existent précisément pour véhiculer un sens ; c'est d'ailleurs le sens que l'on veut exprimer qui déterminera notre choix lexical. Le signifié d'une lexie est une autre caractéristique qui est figée synchroniquement et soumise à la norme prescriptive.

Nous avons vu que trois parties composent une définition : le genre prochain, les différences spécifiques et le typage des actants. Trois transgressions seront donc possibles en ce qui concerne le signifié. Le premier type de transgression sera le non-respect de la composante de sens centrale : le genre prochain. Nous nommerons une telle erreur **impropriété** et, pour des raisons que nous expliquerons plus tard, nous la classerons parmi un autre type de problèmes : les cas d'utilisation d'une lexie fictive (voir II-2.3.4). Les deux erreurs relatives au sens que nous retenons sont donc le **non-respect du typage des actants** et le **non-respect d'une ou de plusieurs différences**

¹⁶ Nous sommes tout à fait conscient qu'il existe différents types de déformation du signifiant, qui correspondent à des phénomènes divers (épenthèse, inversion de phonèmes, etc.). Par contre, étant donné le nombre déjà élevé de problèmes lexicaux que nous présentons dans notre recherche, nous n'avons pas jugé nécessaire d'introduire des distinctions plus fines pour ce type particulier d'erreur.

spécifiques. C'est dans la parole que nous pourrions constater que le sens d'une lexie a été altéré, puisqu'il y aura cooccurrence avec d'autres lexies avec lesquelles elle est sémantiquement incompatible.

La combinatoire sémantique est imposée par le sens des unités lexicales et restreint les possibilités de cooccurrence. Comme nous l'avons mentionné à la section II-1.4, le typage des actants sémantiques dans la définition d'une lexie prédicative oriente le choix des unités lexicales qui pourront remplir le rôle d'argument du prédicat. Par exemple, la phrase *#Il ne faut pas déplacer un blessé pour ne pas l'endommager* présente une incompatibilité sémantique, puisque la structure actancielle d'ENDOMMAGER est *X endommage la chose Y*. Or, la lexie BLESSÉ, qui dénote un être animé (enfin, aussi animé que son état le lui permet!), ne peut jouer le rôle de deuxième argument auprès du prédicat ENDOMMAGER, car ce dernier doit dénoter une chose (ex. *Il ne faut pas déplacer cette statue pour ne pas l'endommager*).

Si le typage des actants peut rendre compte de certaines incompatibilités sémantiques, il faudra recourir à des composantes « internes » du sens des unités lexicales pour en expliquer d'autres : les différences spécifiques. Prenons l'exemple du verbe MINIMISER, qui est défini ainsi par le logiciel *Antidote Prisme* : « Diminuer les proportions, l'importance de (qqch.), dédramatiser. » Ainsi, une phrase telle *#Après s'être fait reprocher ses excès, la gouverneure générale a choisi de minimiser ses dépenses* sera sentie comme problématique, puisque les différences spécifiques du sens de MINIMISER ne sont pas respectées; MINIMISER ne signifie pas seulement 'diminuer Y', mais 'diminuer les proportions, l'importance de Y'. Si le sens est soumis à la norme prescriptive, il faut tout de même comprendre que dans certains contextes humoristiques (notre exemple pourrait d'ailleurs être perçu comme tel) ou poétiques, la transgression du principe de combinatoire sémantique sera acceptable; il reviendra donc au correcteur de juger de l'incompatibilité selon le contexte (cf. norme fonctionnelle).

Les propriétés de combinatoire :

Nous avons vu qu'une des dimensions du syntactique de l'unité lexicale est la **combinatoire grammaticale** (voir II-1.5.1/2), qui regroupe le régime de la lexie et l'ensemble des propriétés grammaticales qui lui sont inhérentes (genre, partie du discours, etc.).

Le régime « correspond aux contraintes syntaxiques attachées à l'emploi d'un terme. Sont ici principalement en cause les constructions prépositionnelles et la transitivité »

(Patris et Vansnick, 1992 : 54). La transitivité ou l'intransitivité, caractéristique inhérente au verbe et figée dans le lexique, doit être respectée si l'on veut que celui-ci fonctionne correctement dans la langue.

Ainsi, nous considérerons comme une erreur lexicale une phrase comme **Il a quitté très tôt*, puisque le caractère transitif de QUITTER n'est pas satisfait. La phrase **Le professeur instruit aux élèves* contient aussi une erreur de régime puisque que le caractère transitif direct d'INSTRUIRE est transgressé. Le choix des prépositions servant à introduire les compléments des verbes, et même des noms, est aussi imposé par chaque unité lexicale. Le choix d'une lexie entraînera automatiquement le choix des prépositions qu'elle régit ; c'est le phénomène de microsyntaxe lexicale. Ce phénomène de microsyntaxe, source de nombreuses erreurs – particulièrement chez les apprenants d'une langue seconde – est difficilement prévisible, et n'est pas encodé de façon systématique dans les dictionnaires (nous savons par contre qu'il existe des ouvrages de référence qui s'intéressent exclusivement à cette question – cf. *Le prépositionnaire*, *Quelle préposition ?*, etc.). Une phrase comme **Il l'a forcé de sortir* contient une erreur de régime, puisque que le verbe FORCER [qqn.] appelle la préposition À. Le verbe PROTESTER (dans le sens de 's'opposer') se construit avec la préposition CONTRE; la phrase **Il proteste la décision du gouvernement* présente donc aussi un problème de régime. Ces deux dimensions, la transitivité et les prépositions régies, constituent des informations statiques stockées à l'intérieur de chaque unité lexicale et sont donc aussi soumises à la norme prescriptive, leur transgression occasionnant des erreurs lexico-syntaxiques.

Nous avons défini le régime comme l'ensemble des contraintes syntaxiques attachées à l'emploi d'un terme, mais d'autres **propriétés de combinatoire grammaticale inhérentes à la lexie** font aussi partie intégrante de sa combinatoire et sont des informations statiques soumises à la norme prescriptive. Le genre nominal et la partie du discours en sont des exemples simples : un nom commun sera soit féminin, soit masculin et un adjectif ne modifiera pas un verbe. Ainsi, **une avion* ou **il marche lent* seront bien entendu considérées comme des erreurs relevant du lexique.

Certaines unités lexicales imposent aussi d'autres types de contraintes grammaticales. Par exemple, de nombreux verbes ou locutions demandent d'être suivis par le subjonctif. Les verbes REGRETTER, CONSENTIR [à], AVOIR HÂTE, CRAINDRE, et les conjonctions de subordination DE PEUR QUE, QUOIQUE, BIEN QUE, AU LIEU QUE sont autant d'unités lexicales qui imposent le mode subjonctif. Une fois de plus, comme cette particularité est

inhérente à de telles lexies, il faudra considérer que sa transgression entraînera une erreur lexicosyntaxique. Il en est de même pour les autres propriétés grammaticales des lexies, par exemple le fait de n'exister qu'à la forme singulier, de ne pas pouvoir se conjuguer à tous les temps, de ne pas admettre la forme passive, etc.

2.3.3 LES DIMENSIONS MOINS « PRESCRIPTIBLES » DU LEXIQUE

Comme nous l'avons mentionné plus haut, l'erreur lexicale ne saurait être définie exclusivement par rapport à la norme prescriptive. En effet, les lois régissant l'emploi du lexique dans la langue demeurent plutôt floues, et ce, même dans la langue écrite.

« Le lexique pose problème parce que n'ont pas encore été découverts ses modes de structuration, ses régularités profondes : on parle depuis longtemps de règles de syntaxe, et non de règles de vocabulaire. [...] La masse des écarts en matière de vocabulaire amène-t-elle les puristes à s'intéresser prioritairement au lexique, mais en même temps la difficulté d'édicter des règles en ce domaine les oblige-t-elle à des remarques ponctuelles : il est possible de pointer une erreur, il est facile de se moquer des gens, on peut même risquer une explication linguistique si on la trouve, mais il est impossible pour rendre compte du système lexical de construire l'équivalent d'une grammaire » (Leeman-Bouix, 1994 : 111, 113).

Même si ce commentaire de Leeman-Bouix nous semble un peu dur en regard des nombreux travaux de lexicologie et de sémantique qui sont quand même arrivés à dégager une certaine structure du lexique, il est vrai que l'élaboration d'une « grammaire » contenant l'ensemble des règles lexicales est une tâche colossale et loin d'être achevée. Par conséquent, peu de règles strictes baliseront l'usage du vocabulaire, et ceci sera particulièrement vrai en ce qui concerne la combinatoire lexicale. Nous pourrions donc toujours condamner aisément un mot-forme dont l'orthographe est incorrecte (par rapport à la norme prescriptive), mais il pourra être beaucoup plus ardu d'expliquer pourquoi telle unité lexicale serait plus appropriée que telle autre. Le lexique est le domaine de la langue qui laisse le plus de latitude, offre le plus de souplesse au locuteur; c'est essentiellement par lui qu'il pourra se permettre un peu d'originalité et de créativité. Si les règles grammaticales sont très strictes et les possibilités syntaxiques plutôt limitées, les possibilités lexicales, elles, sont infinies. Une approche strictement normative (ou prescriptive) en matière de lexique serait une grave erreur, puisqu'elle réduirait la langue à un ensemble fini de possibilités figées. Or, il n'y a pas qu'une « bonne » façon d'exprimer quelque chose. Doit-on conclure

que l'erreur lexicale n'existe pas au-delà des problèmes d'orthographe ou de régime? Que les unités lexicales peuvent être assemblées indifféremment en autant qu'elles respectent les règles syntaxiques? Bien sûr que non.

Le signe linguistique, comme nous l'avons vu plus haut (voir II-1.3.1), comporte trois composantes : le signifié (le sens), le signifiant (la forme) et les propriétés de combinatoire. La plupart des dictionnaires commerciaux ne présentent systématiquement que le signifiant (par ordre alphabétique) et le signifié (par une définition). Les propriétés de combinatoire, quant à elles – hormis la partie du discours et quelques détails concernant le régime –, sont souvent laissées pour compte ou présentées de façon implicite. On constate donc que les aspects du lexique qui pourront être jugés en fonction de la norme prescriptive sont ceux que l'on retrouve dans les ouvrages de référence. C'est à propos des autres aspects de la lexie, ceux faisant partie du syntactique, que la notion de norme lexicale devient plus floue.

Conséquemment, une erreur relevant d'un problème de combinatoire sera plus difficile à repérer, à circonscrire et à expliquer. Dans l'élaboration de notre typologie, nous devons nécessairement tenir compte de cette particularité du lexique et reconnaître humblement qu'il sera impossible, dans certains cas, d'évacuer complètement toute subjectivité lorsque nous qualifierons d'erreur un problème de combinatoire. Par contre, comme il est important pour nous de pouvoir statuer sur le caractère erroné ou non d'un problème lexical relevant du syntactique, nous essaierons, autant que faire se peut, de nous référer à des ouvrages ou de recourir à des techniques de vérification pour réduire cette part de subjectivité malheureusement inévitable. Comme nous adoptons malgré tout une approche prescriptive en qualifiant problèmes lexicaux des usages relevant des dimensions moins prescriptibles du lexique, nous nuancerons en introduisant le terme de **maladresse lexicale** à côté de celui d'*erreur lexicale*, qui semble dénoter une condamnation catégorique (voir III-2.5.3).

La cooccurrence lexicale restreinte est cette propriété qu'ont les unités lexicales d'apparaître ensemble de façon contrainte dans un même énoncé. La connaissance d'une unité lexicale doit donc impliquer celle des cooccurrences, ou collocations, qu'elle contrôle. Ainsi, un étudiant qui prétend connaître le lexème ENQUÊTE devrait savoir que l'on *mène/conduit une enquête*, que cette *enquête porte sur qqch./concerne qqch.*, qu'elle *avance ou piétine*, etc. Toutes ces informations sont essentielles au bon fonctionnement d'une lexie dans la langue ; elles sont difficiles à acquérir de façon systématique et demeurent souvent l'apanage du locuteur natif. Dans les dictionnaires

commerciaux, toutes les informations sur la cooccurrence lexicale sont omises, ou alors introduites de façon implicite dans les exemples. Il n'existe pas, à proprement parler, une référence qui ait autorité en matière de cooccurrences; c'est plutôt l'usage qui figera ou privilégiera certaines cooccurrences. On peut donc affirmer que la norme qui concerne les collocations sera plutôt descriptive, puisque statistique.

Les expressions idiomatiques, ou **locutions**, sont figées et souvent capricieuses dans leur utilisation. Par exemple, la locution FAIRE UNE CROIX [*sur qqch.*] signifie en français 'renoncer'. Il serait donc tout à fait correct de dire : *Le jeune homme avait fait une croix sur son voyage.* Par contre, la forme passive de la locution semble plutôt bizarre : **Une croix avait été faite sur son voyage.* Évidemment, cette particularité de la locution fait en quelque sorte partie de ses propriétés de combinatoire grammaticale. Si nous avons choisi de considérer le caractère figé des locutions dans les dimensions moins prescriptibles du lexique, c'est que les ouvrages de référence ne donnent que très peu d'informations à ce propos et que c'est plutôt l'usage qui nous dictera quelles manipulations sont acceptables pour chaque locution; pourtant, le caractère plus ou moins rigide des locutions doit être respecté pour qu'elles soient utilisées correctement.

En nous inspirant de Schmitt (2000), nous considérons le **registre de langue** comme une connaissance lexicale. Ainsi faut-il savoir à quel niveau de langage ou à quelle variété appartient une unité lexicale que nous désirons utiliser. Le registre de langue doit être adapté à la situation de communication; c'est la norme fonctionnelle qui nous dictera celui à privilégier. À l'écrit particulièrement, le registre de langue doit être homogène et l'usage d'expressions familières ou vulgaires serait perçu comme maladroit (*Le premier ministre a invité ses ministres à une réception suivie d'une [?]bouffe*).

Les québécismes seront traités de la même façon que les autres unités lexicales : non pas comme des régionalismes, mais plutôt comme des unités lexicales de la langue qui doivent être utilisées en harmonie avec le contexte. Nous sommes en effet persuadé qu'il existe un français québécois standard, mais, comme le notent Cajolet-Laganière et Martel (1995 : 12) « le « modèle » auquel renvoie obligatoirement le « bon » usage de la langue écrite [au Québec] n'est pas encore décrit. » Il serait tout à fait aberrant de sanctionner une unité lexicale au Québec sous prétexte qu'elle ne figure pas dans un dictionnaire français de France. « C'est dans tous les domaines de la vie courante et professionnelle que l'on retrouve des écarts entre le français québécois et le français de

référence, c'est-à-dire celui qui est décrit dans les dictionnaires français » (Martel et Cajolet-Laganière, 1996 : 99). Un étudiant québécois doit donc pouvoir utiliser librement l'ensemble du lexique que lui offre sa langue, en respectant les restrictions imposées par le contexte. « En effet, on note que le vocabulaire, en raison même de son caractère explicite, sert de marqueur d'appartenance à une communauté linguistique particulière, de moyen d'affirmation de sa propre identité culturelle [...] » (Martel et Cajolet-Laganière, 1996 : 83). Ainsi, des québécismes comme BABILLARD, CREUX (au sens de 'profond'), MAGASINER, DÉBARBOUILLETTE, SOUVERAINISTE OU POUTINE¹⁷ ne seront pas considérés comme des erreurs. Par contre, une lexie comme ASTHEURE, courante au Québec et parfaitement acceptable lors d'une conversation dans la « vraie vie », serait inappropriée dans une rédaction à l'école ou un article scientifique. Cette unité lexicale appartenant à un registre familier, elle serait considérée comme une maladresse lexicale dans un contexte écrit où le français québécois standard s'impose. Bien sûr, il en serait tout autrement dans un roman québécois de Michel Tremblay, par exemple. C'est donc la norme fonctionnelle qui nous sert dans un tel cas à repérer et à expliquer une erreur lexicale.

Les mélanges de codes. Nous avons vu plus haut que la situation géographique est une des sources de variation de la langue. Il arrive que des mélanges s'effectuent lorsque deux (ou plusieurs) langues évoluent ensemble. Une même langue peut aussi évoluer différemment dans des régions différentes; il s'agit des variantes dialectales. Ces variantes « locales » – qu'elles se manifestent par des lexies nouvelles, des emplois vieillis ou simplement des emprunts tels quels à d'autres langues – sont souvent marquées par rapport à la langue écrite standard. Les anglicismes en français constituent un bon exemple de mélange de codes. Certains de ces emprunts ou de ces variantes dialectales restent dans la langue et finissent par s'intégrer au lexique de la langue standard (par exemple WAGON). Par contre, on tente, particulièrement au Québec, de diminuer l'intrusion de l'anglais dans le français, notamment en bannissant l'usage d'un terme anglais (du moins à l'écrit) lorsqu'une variante française existe. Il existe bien sûr plusieurs types d'anglicismes; le *Dictionnaire des anglicismes Le Colpron* (1999) en définit six : sémantique, lexical, syntaxique, morphologique, phonétique et graphique. Il est intéressant de noter que les auteurs de ce dictionnaire ont regroupé sous les anglicismes syntaxiques autant les usages transgressant les règles

¹⁷ Il est intéressant de noter qu'il existe différents types de québécismes. De Villers (1999) en distingue quatre : les québécismes issus du fonds français (MENTERIE, MOUILLER = 'pleuvoir'), les québécismes de création (CÉGÉPIEN, DÉPANNEUR), les québécismes d'emprunts (CORONER, ACHIGAN) et les québécismes calqués (CRÈME GLACÉE, FIN DE SEMAINE). Pour une typologie exhaustive, voir De Villers (1999).

de la syntaxe du français (**C'est une fille drôle à parler avec.*) que les calques relevant de la microsyntaxe lexicale (**Ils vous ont assigné à cette tâche.*). Bien entendu, nous considérerons seulement les dysfonctionnements relevant de la microsyntaxe dans notre étude des erreurs lexicales.

Deux types de maladroites relatives au lexique n'ont pas été mentionnés jusqu'à maintenant, puisqu'elles ne sont pas inhérentes au lexique même, ne résultent pas d'une « frustration » d'un des aspects de l'unité lexicale. Ces erreurs découlent plutôt de l'utilisation maladroite du lexique dans un texte; elles peuvent par conséquent être difficiles à détecter, à expliquer et à pénaliser. Ces deux sources d'erreurs sont l'**imprécision du vocabulaire** et la **répétition**. Notons d'ailleurs que ces aspects sont toujours évalués de façon qualitative, notamment à l'épreuve uniforme de français (Moffet, 1999). L'imprécision de vocabulaire se manifeste généralement de deux façons (Patris et Vansnick, 1992 : 49) ; l'utilisation de périphrases au lieu du terme propre (ex. *Il se tenait à l'endroit où l'on attend l'autobus.* = *l'arrêt d'autobus*) et l'emploi de « mots passe-partout » – ou termes génériques (ex. *CHOSE*). La répétition peut aussi prendre deux formes; soit la répétition excessive d'une même unité lexicale dans une phrase ou un texte, soit la répétition en écho (ex. *Elle avait dessiné un dessin sur le mur*). Finalement, nous pourrions considérer que le **pléonasme** est une variante sémantique de la répétition, puisqu'il met en cooccurrence deux lexies dont l'une « répète » une composante de sens déjà présente dans l'autre (ex. *chuchoter à voix basse*).

2.3.4 UN AUTRE TYPE DE PROBLÈME LEXICAL : L'UTILISATION D'UNE LEXIE FICTIVE

Nous avons énuméré plus haut différents types de problèmes lexicaux qui découlent, pour la plupart, d'une « altération » d'une des dimensions de l'unité lexicale : signifiant, signifié ou propriétés de combinatoire. Par contre, nous n'avons pas considéré l'utilisation d'une unité lexicale inexistante, qui constituerait bien sûr une erreur lexicale. Nous distinguerons trois cas de figure possibles.

La création d'une nouvelle acception d'un vocable :

La création d'une nouvelle acception d'un vocable, souvent en suivant un patron de polysémie productif, constitue le premier cas d'utilisation d'une lexie fictive. Par exemple, la métonymie et la métaphore constituent deux sources de création de nouvelles acceptions d'un vocable. « Une lexie L_2 est liée de façon métonymique à une lexie L_1 du même vocable si elle dénote un concept lié par une forme de proximité (dans le sens le plus général) au concept dénoté par L_1 . On parle alors de contiguïté de concepts. » (Polguère, 2003 : 166). Par exemple, l'acception du vocable KREMLIN signifiant 'gouvernement russe' [*Le Kremlin a exprimé ses sympathies aux familles des victimes.*] est une métonymie de l'acception de base qui signifie 'forteresse située à Moscou abritant les locaux du gouvernement russe' [*Des milliers de visiteurs envahissent le Kremlin chaque été.*]. Le lien unissant deux lexies dont une est une métaphore de l'autre est plutôt un lien de ressemblance. Par exemple, l'unité lexicale TACHE [*Le petit voisin est une vraie tache.*], plutôt familière et commune en français québécois, signifie 'personne qui impose sa présence, qui importune et dérange constamment.' (Antidote Prisme, 2003 : article TACHE). Il est facile de comprendre que cette acception du vocable TACHE a été créée sur la base d'un lien de ressemblance avec l'unité lexicale qui dénote une marque salissante qui s'imprègne et dont on a peine à se débarrasser.

Certaines métaphores ou métonymies sont lexicalisées, comme les exemples que nous venons de donner, mais il est aussi possible d'en créer de nouvelles. Par contre, certaines de ces créations pourront parfois être perçues comme maladroites ou tout simplement erronées, soit parce que le lien de sens n'est pas assez clair, ou alors parce que le contexte dans lequel elles sont créées ne s'y prête pas vraiment.

D'autres patrons de polysémie existent aussi, et pourront occasionner des erreurs. Les verbes en français regroupent souvent deux lexies dont l'une est le causatif de l'autre. Les lexies CHAUFFER1 [*La sauce chauffe sur le feu.*] et CHAUFFER2 [*Je chauffe la sauce.*] illustrent bien ce phénomène. CHAUFFER1 a la structure actancielle *X chauffe*, alors que CHAUFFER2 a la structure *X chauffe Y* et signifie 'X cause que Y chauffe1'. En appliquant ce patron très fréquent en français à la lexie CLIQUETER [*Ses nombreux bijoux cliquetaient.*], quelqu'un pourrait ainsi créer une nouvelle unité lexicale qui en serait le causatif. Bien entendu, cette lexie n'existe pas en français et la personne qui écrirait **Il cliqueta ses clefs pour attirer son attention* commettrait une erreur lexicale. Ce serait la même chose pour quelqu'un utilisant le verbe pronominal (se) PÉRIMER de façon non pronominale, c'est-à-dire en en modifiant la structure actancielle. Ainsi, la phrase **La baisse d'achalandage*

a périmé les stocks de nourriture présente aussi une acception fautive d'une unité lexicale. La raison pour laquelle nous ne considérons pas ces erreurs comme une altération du signifié est que, en ajoutant un actant sémantique à la lexie, nous créons un nouveau sens, donc une nouvelle unité lexicale. De plus, il est fort probable que la personne qui commettra une telle erreur connaisse tout à fait les lexies CLIQUETER et (se) PÉRIMER à partir de laquelle elle aura créé sa lexie fictive.

L'impropriété :

L'impropriété résulte d'une mauvaise association entre un sens et un signifiant déjà existant. Par exemple, un étudiant qui écrirait la phrase *Mes parents ont toujours été très condescendants* pour exprimer 'mes parents ont toujours été très attentionnés' commettrait une impropriété, puisque la lexie CONDESCENDANT dont le sens est 'attentionné' n'existe pas. Nous avons vu plus haut que le sens d'une lexie peut être altéré soit en ne respectant pas le typage des actants, soit en ne tenant pas compte d'une différence spécifique. Dans le cas de l'impropriété, il ne s'agit pas simplement d'une altération du sens, mais d'une nouvelle association entre une forme et un sens, donc de la création d'une nouvelle unité lexicale. C'est pour cette raison que nous dissociions l'impropriété de la « simple » transgression du sens mentionnée en 2.3.2, et que nous parlons d'utilisation d'une lexie fictive.

La mauvaise dérivation morphologique :

Une **mauvaise dérivation morphologique** peut aussi être à l'origine de l'utilisation d'une lexie inexistante. Par exemple, sur le modèle de ATERRIR ~ ATERRISSAGE, CAPTER ~ CAPTEUR et BANQUE ~ BANQUIER, il serait possible d'obtenir les lexies fautives *BLANCHISSAGE (d'argent), *ADAPTEUR et *BROCANTIER. Nous croyons qu'il faut distinguer le cas de la mauvaise dérivation morphologique de celui du barbarisme présenté plus haut (voir II-2.3.2), puisque le signifiant de la lexie n'a pas été déformé, mais généré à partir d'une autre lexie par une règle de dérivation. Ces erreurs de dérivation morphologique témoignent, comme le mentionnent Patris et Vansnick (1992 : 51), d'une certaine compétence linguistique, parce qu'elles « ont été créé[e]s par l'application de règles fonctionnant tout à fait pertinemment dans la langue ».

Une dernière remarque s'impose à propos des trois premiers cas d'utilisation d'une lexie fictive que nous venons de présenter. Nous avons caractérisé les types d'erreurs

lexicales mentionnés en II-2.3.2 et II-2.3.3 par leur description structurale, sans chercher à en décrire la cause. C'est d'ailleurs l'approche que nous privilégierons aussi souvent que possible, et nous justifierons ce choix dans notre méthodologie (voir III-2.5.1). Par contre, dans le cas de l'utilisation d'une unité lexicale inexistante, nous nous voyons contraint de spéculer sur la cause de l'apparition de cette lexie, puisqu'il est évident que les sources sont multiples. Nous sommes conscient qu'en nous éloignant de la simple description structurale du dysfonctionnement, nous nous exposons à plus de subjectivité et donc à plus d'occasions d'interprétations fausses. Prenons par exemple la phrase *Le roi vivait dans la **luxure*** écrite pour exprimer 'le roi vivait dans la richesse'. Nous pouvons considérer qu'il s'agit d'une impropriété, puisque le signifiant *luxure* est ici associé au signifié 'richesse', et qu'aucune lexie du français ne présente cette association signifiant ~ signifié ; ainsi, nous supposons qu'il existe dans le vocabulaire de l'auteur de cette phrase une lexie LUXURE qui n'existe pas dans le lexique français. Par une autre interprétation, nous pourrions plutôt opter pour la mauvaise dérivation morphologique : à partir de la lexie LUXE qu'elle maîtrise, la personne a dérivé la fausse lexie *LUXURE (= 'luxe'), sur le modèle de JOINT ~ JOINTURE ou de ÉCRIT ~ ÉCRITURE. Finalement, nous pouvons spéculer que l'utilisation de LUXURE dans la phrase erronée relève des deux types d'erreurs, puisque cette lexie fictive « existe » vraiment dans le vocabulaire de la personne – ce qui en fait une impropriété –, mais que son existence est en quelque sorte validée par une parenté morphologique avec la lexie LUXE qui laisse croire à une dérivation morphologique « normale ». Certaines erreurs associées aux trois premiers cas d'utilisation d'une lexie fictive pourront donc être caractérisées comme relevant de plusieurs types d'erreurs à la fois.

CONCLUSION

Nous avons donc terminé notre caractérisation des types d'erreurs lexicales en fonction des différentes dimensions de la lexie. L'aboutissement de cette exploration sera une typologie des problèmes lexicaux que nous présenterons dans notre méthodologie (voir III-2.5).

3. LES STRATÉGIES DE RÉOLUTION DE PROBLÈMES LEXICAUX

Le deuxième objectif de la présente recherche vise à décrire les stratégies de résolution de problèmes lexicaux des étudiants universitaires, aussi est-il important de bien définir ce que nous entendons par les termes *problème* et *stratégies de résolution de problèmes*. La section qui suit caractérisera en un premier lieu ces deux notions avant d'expliquer comment elles s'appliquent au lexique.

3.1 CARACTÉRISATION DE LA NOTION DE PROBLÈME

« Glover, Ronning et Bruning (1990) précisent qu'un problème existe quand il y a un ou des obstacles qui séparent l'état initial, c'est-à-dire les données de la situation, du but désiré, c'est-à-dire la question à laquelle une personne doit répondre ou la tâche qu'elle doit réaliser » (Tardif, 1992 : 235). Selon Tardif (1992), un problème à résoudre possède quatre caractéristiques :

- 1- présence de données initiales;
- 2- présence d'un état désiré;
- 3- présence de contraintes;
- 4- nécessité de la recherche d'une suite d'opérateurs.

Par exemple, un étudiant en situation de production de texte se retrouverait face à un problème s'il butait sur l'orthographe d'un mot-forme. La présence de données initiales serait le choix de l'unité lexicale en question. L'état désiré, ou le but, serait de pouvoir utiliser cette lexie avec l'orthographe correcte. La présence de contraintes pourrait se traduire par l'impossibilité de recourir au dictionnaire. Résulterait de cette situation la nécessité de faire une recherche cognitive active et de trouver un moyen de parvenir à une solution (par exemple se remémorer un dérivé morphologique qui pourrait l'éclairer sur la finale du mot-forme). Nous considérerons donc qu'il y a problème chaque fois qu'un scripteur hésite sur ce qu'il doit écrire ou qu'un réviseur s'interroge sur une trace graphique.

3.2 LA RÉOLUTION DE PROBLÈMES

Pour passer de l'état initial au but désiré, la personne devra nécessairement franchir différentes étapes qui constituent la *résolution du problème*. Andre (1986) dénombre six étapes dans la résolution de problème, plus une septième facultative :

Tableau I. *Étapes de la résolution d'un problème (Andre, 1986, cité dans Tardif, 1992 : 238)*

Phase de représentation du problème :
<ul style="list-style-type: none"> • Reconnaissance du problème. • Description du problème. • Analyse du problème.
Phase de solution du problème :
<ul style="list-style-type: none"> • Génération d'un scénario de résolution. • Évaluation de l'efficacité des solutions privilégiées. • Mise en application de la solution retenue. • (Mise en application de nouvelles solutions au besoin.)

Phase de représentation du problème :

La **reconnaissance du problème** est la prise de conscience de l'écart entre l'état initial et le but désiré.

- *Doit-on écrire « ce que j'ai subi ce matin » ou « ce que j'ai subi ce matin » ?*

La **description du problème** vise à déterminer les différents paramètres du problème.

- *L'état initial est le syntagme ci-haut, le but désiré est de savoir si « subi » prend un « t » ou non et la contrainte est que je n'ai pas droit aux ouvrages de référence.*

L'**analyse du problème** consiste à définir ce qui est connu ou inconnu à propos de ce problème.

- *« Subi » est ici un participe passé employé avec l'auxiliaire « avoir » ; il doit donc s'accorder avec le complément direct s'il est placé avant. Le complément direct de « subi » est « que » placé avant mis pour « ce ». Donc « subi » devra s'accorder au masculin singulier. Maintenant, est-ce que le participe passé du verbe « subir » prend un « t » en finale ou non ?*

Phase de résolution du problème :

La **génération d'un scénario de résolution de problème** est l'étape durant laquelle la personne confrontée à un problème décide d'une ou de plusieurs solutions pour parvenir au but désiré.

- *Je pourrais conjuguer d'autres verbes du même groupe, c'est-à-dire les verbes en « -ir », par exemple « finir », pour voir quelle est la finale de leur participe passé. Ou encore, je pourrais construire une phrase dans laquelle le participe passé s'accorderait au féminin pour voir si le participe passé sera « subie » ou « subite ».*

La prochaine étape est l'**évaluation de l'efficacité des solutions privilégiées** dans le cas où plusieurs solutions ont été envisagées.

- *Ça ne sert à rien de comparer avec « finir », car je ne sais pas non plus si on doit écrire « le livre que j'ai finit ce matin » ou « le livre que j'ai fini ce matin ». Je vais plutôt accorder le participe passé au féminin.*

Suit la **mise en application de la solution retenue**.

- *« La honte que j'ai subie ce matin » ou « la honte que j'ai subite ce matin » ? Les deux ont l'air de fonctionner... « subie », ça sonne bien... mais « subite », ça se dit aussi. On dit « une mort subite »...*

Au besoin, la personne pourra mettre en **application de nouvelles solutions** si la solution choisie plus tôt n'est pas satisfaisante.

- *Si je mets au féminin avec « finir », est-ce que les deux fonctionnent? « La page que j'ai finie » ou « la page que j'ai finite » ? Ouache! « finite », ça existe pas! C'est « finie » c'est sûr. Alors j'imagine que c'est aussi « la honte que j'ai subie ». Donc, « subi » prendra pas de « t » à la fin.*

Voilà donc l'ensemble des étapes par lesquelles pourra passer une personne confrontée à un problème d'une telle nature. Nous appellerons **stratégies de résolution de problèmes** l'ensemble des processus mis en œuvre au cours des différentes étapes de la résolution du problème.

Notons par ailleurs que la résolution de problème en six ou sept étapes ne s'applique pas qu'au domaine du français écrit, mais à l'ensemble des problèmes, de quelque ordre qu'ils soient (mathématique, social, logique, etc.). Gagné (1985) souligne que c'est la phase de représentation du problème qui est cruciale, puisque c'est à ce moment que la personne détermine les connaissances auxquelles elle devra faire appel pour résoudre le problème. Les experts d'un domaine consacrent généralement

beaucoup de temps à la représentation du problème avant d'envisager des solutions. Les novices, de leur côté, tendent à appliquer aveuglément des solutions sans nécessairement avoir une représentation complète du problème.

3.3 TRANSPOSITION DE CES CONCEPTS AU LEXIQUE

Voyons maintenant, à la lumière des informations que nous venons de présenter, de quelle façon nous pouvons définir ce qu'est un *problème lexical* (II-3.3.1). Le point II-3.3.2 vise à clarifier une ambiguïté entourant l'expression *problème lexical*. Dans le point II-3.3.3, nous justifierons notre choix de réduire le nombre d'étapes pour la résolution de problème appliquée au lexique et nous expliquerons ce que nous entendons par *stratégies de résolution de problèmes lexicaux*.

3.3.1 LE PROBLÈME LEXICAL

Un *problème lexical* surviendra chaque fois qu'un scripteur ou qu'un réviseur se questionnera à propos d'une dimension relevant du lexique en cours de tâche. Compte tenu des différentes dimensions du lexique (voir II-1.3.1), nous pouvons déduire qu'un problème lexical pourra être d'ordre graphique, sémantique, grammatical, contextuel, ou alors découler d'une anomalie par rapport à l'usage courant. Évidemment, une personne pourra se trouver face à un problème lexical sans savoir que ses questionnements concernent un aspect du lexique.

3.3.2 REMARQUES TERMINOLOGIQUES

Nous désirons maintenant souligner que l'expression *problème lexical* dénote, dans le cadre de notre recherche, deux concepts différents. Cette polysémie peut paraître fâcheuse à première vue, mais ne devrait pas compromettre la clarté de notre exposé. Nous venons de présenter, au point II-3.3.1, ce qu'est un problème lexical dans une perspective de psychologie cognitive : il s'agit de tout questionnement à propos d'une dimension lexicale. C'est dans ce sens qu'il faut entendre la deuxième occurrence de *problèmes lexicaux* figurant dans le titre du présent mémoire : *Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes lexicaux et analyse des stratégies*

de résolution de problèmes lexicaux. Le fait que l'expression *stratégie de résolution de problèmes* soit consacrée dans la recherche en didactique et que nous désirions faire appel à ce concept dans notre travail nous a obligé à retenir aussi le terme *problème*.

Par ailleurs, nous parlons de *problèmes lexicaux*, notamment au chapitre IV, pour désigner l'ensemble des dysfonctionnements lexicaux (erreurs et maladroites, voir III-2.5.3) observés dans les copies d'étudiants. C'est d'ailleurs à cette acception que correspond la première occurrence de *problèmes lexicaux* dans le titre de notre mémoire : *Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes lexicaux et analyse des stratégies de résolution de problèmes lexicaux*.

Nous avons donc choisi de conserver l'expression *problème lexical* pour renvoyer à des concepts différents dans notre travail, car nous croyons que c'est celle qui les dénote avec le plus d'éloquence. De plus, nous avons vu à la section II-1.3.5 que la polysémie est un phénomène naturel et courant dans la langue. Nous nous fions donc au bon jugement de nos lecteurs et sommes persuadé qu'ils sauront déterminer sans problème (!) à quelle acception de *problème lexical* nous faisons référence!

3.3.3 LES STRATÉGIES DE RÉOLUTION DE PROBLÈMES LEXICAUX

La résolution de problèmes en six (ou sept) étapes proposée par Andre (1986) s'applique très bien à des problèmes complexes. Par exemple, l'écriture d'un texte est une tâche cognitive très complexe qui nécessitera qu'on passe par toutes ces étapes afin de bien la réaliser.

Les problèmes lexicaux, quant à eux, constituent des problèmes beaucoup plus ciblés, locaux, et par conséquent moins complexes. L'écriture d'un texte comportera d'ailleurs une multitude de problèmes lexicaux. Le problème lexical se situe donc à un autre niveau; nous pourrions parler de « microproblème ».

Les étapes de la résolution d'un problème lexical sont moins bien définies que pour les problèmes complexes, et certaines des étapes proposées par Andre (1986) semblent se chevaucher. Nous simplifierons donc le processus de résolution de problème à 4 étapes.

Tableau II. Étapes de la résolution d'un problème lexical

Phase de représentation du problème :
<ul style="list-style-type: none"> • Repérage du problème • Analyse du problème
Phase de solution du problème :
<ul style="list-style-type: none"> • Génération, application et analyse d'une ou de plusieurs solution(s) • Prise de décision

Nous avons donc fusionné les étapes de description et d'analyse du problème proposées par Andre (1986) ainsi que les étapes de génération et d'évaluation du ou des scénario(s) de résolution. La mise en application de la solution de Andre (1986) correspond à notre prise de décision.

Prenons par exemple le cas d'un étudiant qui écrit la phrase *Il prendra part dans la compétition du mois prochain*. Il s'interroge sur la préposition à utiliser avec le verbe PRENDRE PART.

La première étape est le **repérage du problème**.

- « *Prendre part dans la compétition* », ça sonne mal.

Suit l'**analyse du problème**.

- *Qu'est-ce qui fonctionne pas... « Prendre part », ça se dit me semble... On dit-tu « prendre part dans une compétition » ou « prendre part à la compétition » ? Ouais, j pense que c'est le « dans » qui est bizarre.*

La troisième étape consiste à **générer, appliquer et analyser une solution**.

- *J'veais chercher dans le dictionnaire. Y faut-tu chercher dans « compétition » ou dans « prendre »? Compétition... [trouve l'article de COMPÉTITION] « Recherche simultanée par deux ou plusieurs personnes d'un même avantage... Il ne supporte pas la compétition. Compétition entre partis politiques. Sortir vainqueur d'une compétition. Être en compétition... » Ah! voilà : « Compétition sportive, épreuve ; championnat... Sport de compétition. Esprit de compétition. » Non, ils le disent pas... Peut-être si je cherche « prendre ». [trouve l'article de PRENDRE] Wow, c'est ben trop long! Attends, j'veais regarder « part » à place. [trouve l'article de PART] Ouais, c'est long aussi... « Ce qu'une personne possède ou acquiert en propre. » Non... Ah ! Tiens ! « Être pour qqch. dans. Un acte blablabla —PRENDRE PART À : jouer volontairement un rôle dans (une affaire). intervenir, participer. Prendre part à un travail, aux frais. Aider, contribuer. Prendre part à une manifestation. Il a pris une part très active à l'élaboration de ce projet. » Bon, enfin! Faque on dit « prendre part À une manifestation ».*

Finalement, le scénario de résolution mis en oeuvre devrait permettre une **prise de décision**.

- *Ça doit être la même chose pour « compétition », alors je vais écrire « Il prendra part à la compétition du mois prochain. » Ça sonne ben mieux de même!*

Les *stratégies de résolution de problèmes lexicaux* seront donc l'ensemble des processus mis en oeuvre au cours des différentes étapes de la résolution d'un problème lexical.

Le terme STRATÉGIE peut connoter un ensemble d'actions réfléchies et bien coordonnées. Dans le cadre de cette recherche, il faut plutôt l'entendre comme un ensemble de processus plus ou moins sophistiqués mis en branle de façon plus ou moins consciente par la personne confrontée à un problème lexical. Il se peut donc que les stratégies soient réduites à leur plus simple expression; la résolution d'un microproblème pourra entraîner des microstratégies! Par exemple, le simple fait de porter un jugement de locuteur sur le choix d'une unité lexicale constituera pour nous une stratégie, au même titre que la recherche dans le dictionnaire ou l'analyse grammaticale d'une phrase. Bien entendu, il pourra être intéressant de constater le degré de sophistication des stratégies déployées.

Une remarque similaire vaut pour les étapes de la résolution d'un problème lexical. En effet, il se peut là aussi qu'une étape se limite à un nombre très restreint d'actions. Par exemple, nous pourrions considérer que le simple fait de proposer une unité lexicale en remplacement d'une autre sans justification constitue l'étape de génération, d'application et d'analyse d'une solution. Bien entendu, nous tiendrons compte dans notre analyse du fait que certaines solutions aux problèmes sont appuyées, alors que d'autres semblent gratuites.

CONCLUSION

En résumé, nous parlerons de *problème lexical* chaque fois qu'un scripteur ou un réviseur se questionnera à propos d'une des dimensions du lexique. Évidemment, il n'est pas nécessaire que la personne confrontée au problème lexical soit consciente que ses questionnements sont relatifs au lexique. Nous considérerons quatre étapes dans la résolution de problèmes lexicaux, et l'ensemble des processus déployés pour régler un problème lexical seront appelés *stratégies de résolution de problèmes lexicaux*.

4. LE MÉTALANGAGE

Nous avons vu dans la problématique (voir I-1) que deux recherches s'étant intéressées aux stratégies de résolution de problèmes linguistiques d'étudiants universitaires en étaient arrivées à la conclusion que les étudiants ne maîtrisent pas le métalangage. Comme nous désirons observer la présence ou l'absence d'éléments du métalangage dans les stratégies de résolution de problèmes appliquées au lexique, nous définirons dans la section qui suit ce que nous entendons par *métalangage*.

4.1 REMARQUES TERMINOLOGIQUES

Il s'impose de justifier le choix du terme *métalangage* pour désigner le sous-ensemble des lexies d'une langue permettant de tenir un discours à propos de la langue. En effet, il serait plus cohérent de parler de *langue* et de *métalangue*. Par contre, le terme *métalangage* semble plus répandu en didactique et, pour cette raison, c'est celui que nous retiendrons dans la présente recherche.

Mais le vocable MÉTALANGAGE est lui-même ambigu. Par exemple, Rey-Debove (1997) définit ainsi le métalangage, qu'elle considère comme un synonyme de MÉTALANGUE :

« On appelle *métalangue* (ou *métalangage*) tout discours, oral ou écrit, qui a pour thème le langage, c'est-à-dire dont le signifié dénotatif est le langage lui-même et ce qu'on en dit. Ce discours (qui s'oppose au langage primaire) met en œuvre un lexique métalinguistique [...] dont une partie, la terminologie linguistique, appartient au langage des spécialistes. » (Rey-Debove, 1997 : 344)

Pour sa part, Gombert (1990) désigne comme synonymes les termes *métalangage* et *activités métalinguistiques*, qu'il définit comme « un sous-domaine de la métacognition qui concerne le langage et son utilisation, autrement dit comprenant; 1- les activités de réflexion sur le langage et son utilisation, 2- les capacités du sujet à contrôler et à planifier ses propres processus de traitement linguistique » (Gombert, 1990 : 27). Pour désigner cette acception de MÉTALANGAGE, nous adopterons la position de Cicurel (1985), qui privilégie le terme d'*activité métalinguistique*. « En effet, plutôt que de parler de métalangage qui risque d'être perçu comme un langage formel, spécialisé, nous avons préféré le terme d'*activité métalinguistique*, car celle-

ci peut en effet s'exercer sans la présence d'un lexique autre que celui de la langue « ordinaire ». » (Cicurel, 1985 : 8).

Nous choisissons donc d'employer le terme *activité métalinguistique* pour désigner toute activité de réflexion sur la langue. Le terme *métalangage* dénotera quant à lui le système de signes conventionnels utilisés pour parler de la langue : le métalexique, constitué du lexique propre à la terminologie linguistique et des autonymes¹⁸.

4.2 LE MÉTALANGAGE ET L'ACTIVITÉ MÉTALINGUISTIQUE

La langue est particulière en ce sens qu'il est impossible de la décrire sans l'utiliser (réflexivité) ; il y a donc risque de confusion entre l'objet d'étude et l'étude elle-même. Par conséquent, il sera important d'utiliser un langage spécialisé qui permettra de se distancier de la langue « courante » pour pouvoir désambiguïser le discours à son sujet : le métalangage. « Ce formidable outil permet [...] de rendre compte de la langue prise comme objet de référence, et fonde la terminologie linguistique » (Petiot, 2000 : 121). Des termes comme *sujet*, *verbe*, *pronom* et *subordonnée relative* sont tous des éléments du lexique du métalangage. Le métalangage propre au lexique contient des termes comme *lexie*, *acception*, *polysémie*, *homonymie*, *synonymie*, *sens*, etc.

Il y aura donc activité métalinguistique chaque fois qu'il y aura recours au métalangage pour parler de la langue. Par contre, l'activité métalinguistique¹⁹ ne se limite pas à la seule utilisation d'éléments de la métalangue. En effet, à l'instar de Jakobson (1963), nous qualifierons d'activité métalinguistique toute activité linguistique portant sur le langage, qu'elle utilise ou non des termes de métalangage.

¹⁸ « Les autonymes sont [...] des noms de signes ou de séquences langagières déterminées, qui sont forgé en discours au fut (sic) et à mesure des besoins, et qui forment, avec les mots métalinguistiques codés, le lexique de base du métalangage (métalexique). » (Rey-Debove, 1997 : 345) Par exemple, la phrase *Toujours prend toujours un « s »* contient une occurrence autonymique de TOUJOURS (la première) et une occurrence « typique » de TOUJOURS (la deuxième).

¹⁹ Certains auteurs préfèrent le terme *activité métalangagière* à celui d'*activité métalinguistique*, parce qu'il le trouvent plus « englobant » en ce sens qu'il traduit mieux le fait que l'activité peut ne pas porter exclusivement sur les dimensions strictement linguistiques (sémantique, syntaxe, etc.), mais aussi sur les « contraintes situationnelles et interactives et les contraintes liées à la gestion du discours dans sa globalité » (Dolz, 1998 : 14). Pour une argumentation complète, voir Dolz (1998).

4.3 L'ACQUISITION DU MÉTALANGAGE : UNE FINALITÉ DIDACTIQUE?

Nous présentons plus haut le métalangage comme un « outil précieux » pour réfléchir et verbaliser à propos de la langue, notamment au niveau du français écrit ; quelqu'un maîtrisant les termes du métalangage serait donc plus à même de diagnostiquer et d'expliquer des problèmes linguistiques. Peut-on alors affirmer que la maîtrise du métalangage favorise la maîtrise du code écrit ? Et à l'inverse, doit-on conclure qu'une bonne maîtrise du code implique nécessairement une maîtrise du métalangage ?

La corrélation entre la maîtrise du code écrit et le recours au métalangage ne semble pourtant pas facile à démontrer. Par exemple, la recherche de Lefrançois (2004) révèle qu'en situation de production de texte, les étudiants forts sont ceux qui font le moins appel au métalangage et conclut que les forts résolvent les problèmes de français écrit avec peu ou pas de métalangage. Force est donc de constater que la maîtrise du code n'implique pas nécessairement une bonne connaissance du métalangage. Ceci rejoint l'idée mentionnée plus haut selon laquelle une activité métalinguistique peut tout à fait être exempte de recours au métalangage.

Doit-on alors considérer que l'acquisition de termes du métalangage à l'école ne devrait pas constituer un objectif didactique ? À cette question, Lefrançois et Montesinos-Gelet (à paraître) proposent une réponse parfaitement en accord avec notre position.

« Acquérir le métalangage couramment utilisé pour décrire la langue (les termes d'*adverbe*, de *sujet*, d'*invariabilité*, par exemple) n'est pas une fin en soi : l'objectif de l'enseignement standard du code linguistique demeure l'utilisation par les scripteurs d'une langue conforme aux normes en vigueur, quels que soient les moyens utilisés pour y parvenir. [...] Les manipulations sur la langue devraient, à notre avis, précéder l'enseignement du métalangage : en effet, comment et pourquoi donner un nom à ce qu'on ne comprend pas ? Cependant, une fois cette compréhension installée, nous croyons qu'il devient impératif de coiffer les concepts du métalangage adéquat » (Lefrançois et Montesinos-Gelet, à paraître : 6).

CONCLUSION

En résumé, nous pourrions dire que notre deuxième objectif de recherche tentera de décrire l'activité métalinguistique propre à la résolution de problèmes lexicaux. Une partie de cette description consistera à mettre en lumière les éléments du métalangage utilisés en situation de résolution de problèmes lexicaux et nous permettra de constater si les étudiants connaissent et maîtrisent un certain métalangage lexical.

III. MÉTHODOLOGIE

1. RAPPEL DES OBJECTIFS

Avant de présenter notre méthodologie, il est important de rappeler les objectifs de la présente recherche. Il s'agit d'une recherche descriptive comportant deux objectifs :

1. décrire les types de problèmes lexicaux d'étudiants universitaires soumis à une activité de rédaction de texte et établir la fréquence de chaque type;
2. décrire les stratégies de résolution de problèmes lexicaux d'étudiants universitaires, notamment en fonction des types de problèmes rencontrés.

De façon générale, notre méthodologie se divise elle aussi en deux grandes étapes. La première, plutôt quantitative, vise à dresser un tableau des différents types de problèmes lexicaux observés dans des textes d'étudiants. La deuxième, plutôt qualitative – et exploratoire, le nombre limité de sujets ne nous permettant pas de prétendre à un état des lieux exhaustif – s'intéresse plutôt à décrire les stratégies de résolution de problèmes lexicaux mises en œuvre.

2. OBJECTIF 1 : DESCRIPTION DES TYPES DE PROBLÈMES LEXICAUX OBSERVÉS

C'est par l'étude d'un corpus de 103 textes produits par des étudiants universitaires que nous avons répondu à notre premier objectif. L'analyse de ces textes nous a permis de constater quels sont les types de problèmes lexicaux des étudiants et quelle est la fréquence de chaque type.

2.1 CORPUS D'ANALYSE ET ÉCHANTILLON

Le corpus de textes analysé afin d'observer les problèmes lexicaux des étudiants était constitué de rédactions faites par des étudiants inscrits au cours *Français écrit pour futurs enseignants* (DID1010A)²⁰.

²⁰ Conformément aux exigences du Comité d'éthique de la Faculté des sciences de l'éducation, nous avons fait remplir un formulaire de consentement aux 103 étudiants dont les textes ont été retenus pour

L'ensemble des étudiants inscrits à ce cours sont en première année de baccalauréat à la Faculté des sciences de l'éducation de l'Université de Montréal. Il font partie des différents programmes offerts à la faculté : éducation préscolaire et enseignement primaire, enseignement de l'éducation physique et santé, enseignement de l'univers social au secondaire, enseignement des mathématiques au secondaire, enseignement des sciences et des technologies au secondaire, enseignement du français au secondaire, enseignement du français langue seconde, et enseignement en adaptation scolaire. Nous n'avons retenu pour notre recherche que les textes écrits par des étudiants ayant été scolarisés en français tout au long de leurs études. Notre corpus regroupait 103 textes.

Si l'on se fie aux résultats obtenus au Test de français Laval-Montréal (TFLM), nous pouvons affirmer que les étudiants présents dans notre échantillon se situent dans la moyenne en ce qui concerne leur maîtrise du français écrit. Ce test est imposé à tous les étudiants désirant s'inscrire dans un programme à la Faculté des sciences de l'éducation, et permet de s'assurer qu'ils maîtrisent le français écrit à un certain niveau. La note de passage du TFLM est fixée à 75 % pour la moyenne des deux sections qui le composent : grammaire et vocabulaire. Tous les étudiants obtenant un résultat inférieur à 75% doivent suivre un cours de mise à niveau en français écrit, et leur inscription en troisième année de baccalauréat est conditionnelle à la réussite de ce cours. Pour les étudiants ayant obtenu un résultat inférieur à 75%, mais égal ou supérieur à 60%, le cours de mise à niveau imposé par l'université est le cours *Français écrit pour futurs enseignants (DID1010A)*.

Certaines raisons ont justifié le choix de notre échantillon. La première raison repose sur l'accès à l'information. En effet, nous voulions disposer d'un corpus assez imposant constitué de textes écrits dans un cadre scolaire afin de s'assurer d'une certaine homogénéité dans le style et la longueur des textes produits. Notre travail comme chargé de travaux pratiques pour le cours *Français écrit pour futurs enseignants (DID1010A)* nous permettant un contact privilégié avec les chargés de cours responsables, nous avons pu avoir facilement accès aux rédactions des étudiants. Ainsi nous fut-il possible de recueillir en peu de temps plusieurs dizaines de textes produits dans le cadre d'un cours universitaire de première année.

Deuxièmement, nous voulions aussi un échantillon qui soit représentatif du niveau en français écrit de la population étudiante en général. Considérant le fait que près des

trois quarts des étudiants obtiennent une note inférieure à 75% au TFLM, nous croyons que notre échantillon est représentatif de la population inscrite à la Faculté des sciences de l'éducation. L'ensemble de nos sujets étant de futurs enseignants, il est bien sûr possible (bien que nous en doutions) que leur compétence en français écrit soit supérieure à celle des étudiants inscrits à d'autres départements et facultés de l'université. Par contre, comme nous excluons de notre échantillon l'ensemble des étudiants ayant obtenu 75% ou plus au TFLM, et que nos sujets, comme leurs collègues des autres facultés et départements, arrivent directement du cégep, il y a lieu de supposer que notre échantillon est assez représentatif de la moyenne des étudiants inscrits en première année à l'université.

2.2 LES TEXTES DU CORPUS

Les textes constituant notre corpus ont été écrits à la maison par les étudiants du cours *Français écrit pour futurs enseignants (DID1010A)*. Ils ont été produits en tout début d'année, lors des premiers cours du trimestre d'automne. Le sujet de ces rédactions formatives était imposé par les responsables des cours; quarante (40) textes portaient sur la discipline en milieu scolaire, trente-quatre (34) étaient des résumés de livres ou de films et vingt-neuf (29) décrivaient un endroit cher à l'auteur durant son enfance. Les textes étaient d'une longueur de 250 mots et, comme la rédaction en a été faite à la maison, le recours aux ouvrages de référence (et aux correcteurs automatiques !) était permis. Comme ces rédactions étaient faites dans un contexte scolaire, nous pouvions nous attendre à ce qu'elles soient écrites dans un français standard ou soutenu.

2.3 LA CORRECTION DES TEXTES

La correction de textes a été effectuée durant l'automne 2004 et l'hiver 2005. Chaque texte a d'abord été photocopié avant d'être corrigé par le responsable du cours. Ce dernier devait relever l'ensemble des erreurs de français présentes dans le texte et mentionner le type d'erreur par un code dans la marge (ex. « orth » pour orthographe lexicale, « gram » pour problèmes d'orthographe grammaticale, « synt » pour problèmes de syntaxe, « voc » pour problèmes de lexique, etc.). Une copie de la correction a ensuite été faite avant que la rédaction ne soit remise à l'étudiant.

Nous nous sommes servi des copies vierges pour effectuer notre propre correction des textes; seulement les problèmes lexicaux ont été relevés. Nous avons ensuite pu vérifier notre correction en la comparant à celle faite par les responsables des cours. Nous avons d'abord prévu établir un accord interjuges avec cette double correction, mais l'idée a dû être abandonnée pour trois raisons. Tout d'abord, comme notre correction visait principalement à découvrir des problèmes lexicaux, nous en avons relevé plus que les responsables de cours qui devaient aussi s'intéresser à tous les autres types d'erreurs et qui disposaient d'un temps restreint pour la correction (une semaine). De plus, nous avons cherché l'ensemble des problèmes lexicaux (erreurs et maladroites, voir section III-2.5.3) présents dans les textes, alors que les autres correcteurs n'ont souligné que les dysfonctionnements considérés comme de « vraies » erreurs. Ensuite, certaines erreurs que nous considérons comme lexicales étaient associées à des catégories différentes par les responsables de cours (notamment « synt » et « gram ») ; cette divergence de points de vue sur certains types d'erreurs a confirmé nos affirmations de la section II-2.3.1, à savoir que certains problèmes que nous considérons de nature lexicale sont plutôt associés à des problèmes de grammaire (notamment de syntaxe) par les enseignants. Bref, plutôt que de nous servir à vraiment valider notre propre correction par un accord interjuges, les copies des corrections effectuées par les responsables des cours ont constitué d'excellents guides pour nous assurer que nous n'oublions pas d'erreurs.

2.4 LES SOURCES DE RÉFÉRENCE

Cinq sources de références ont été retenues pour la présente recherche. Elles nous ont aidé à trancher sur le statut des problèmes lexicaux dans les cas litigieux. Les sources de référence sont les suivantes :

- 1- *Le Petit Robert* : Cet incontournable semble être perçu, autant par les linguistes que par les enseignants, comme la référence parmi les dictionnaires grand public. Avec une nomenclature de 60 000 mots, *Le Petit Robert* constitue un ouvrage de référence précieux relativement aux définitions, à la polysémie et aux exemples proposés. Bien entendu, cet ouvrage décrit principalement le français de France et le lexique du français québécois en est pratiquement absent.
- 2- *Le Multidictionnaire de la langue française, quatrième édition* : Cet ouvrage connaît un grand succès au Québec depuis quelques années déjà. D'une nomenclature plus restreinte que *Le Petit Robert*, le *Multidictionnaire* contient de nombreuses informations sur les difficultés

orthographiques, les problèmes grammaticaux, les anglicismes, les québécoismes, etc. Très normatif, il propose des solutions à une foule de problèmes que peut rencontrer un scripteur. Par contre, il s'avère beaucoup moins riche que *Le Petit Robert* en ce qui concerne les définitions et les exemples.

- 3- ***Antidote Prisme v.5*** : Ce logiciel québécois contient, en plus d'un dictionnaire « traditionnel », un dictionnaire des synonymes, une grammaire, un correcteur et un conjugeur. Moins normatif que le *Multidictionnaire*, il s'avère assez exhaustif en ce qui a trait au français québécois et présente la particularité d'énumérer systématiquement pour chaque mot-vedette l'ensemble des locutions qui le contiennent, ainsi qu'une définition de ces locutions. Le logiciel offre aussi une multitude de fonctionnalités très intéressantes d'aide à la correction de textes. Comme aucun ouvrage ne semble aujourd'hui faire autorité en matière de français québécois, nous avons cru bon introduire deux ouvrages d'ici afin de disposer de plusieurs points de vue.
- 4- ***Le dictionnaire des cooccurrences*** : Peu connu, cet ouvrage québécois regroupe, pour environ 4200 termes du vocabulaire courant, une liste de cooccurrences recueillies au fil des ans et des lectures par Jacques Beauchesne. Très pratique pour un utilisateur averti, ce dictionnaire a le défaut de manquer de structure dans la présentation. En effet, les cooccurents sont donnés en vrac, sans référence systématique au sens qu'ils véhiculent. De plus, si la plupart des cooccurrences fréquentes y apparaissent, un certain nombre de collocations discutables s'y retrouvent aussi. Néanmoins, pour le scripteur sachant faire preuve de discernement, cet ouvrage peut s'avérer un bon guide dans la recherche du mot juste.
- 5- **Le Web** : Pour les aspects du lexique plutôt consacrés par l'usage (les collocations et la rigidité d'emploi des locutions), nous nous sommes réservé le droit de faire une recherche sur le Web. Beaucoup de recherches lexicographiques se font à l'aide de corpus (voir Stubbs, 2001) et le Web, de par son volume, constitue un corpus inestimable pour étudier la langue. En tapant une expression dans le fureteur *Google*, il est possible de constater sa fréquence par le nombre d'occurrences. Bien sûr, une telle méthode demande de faire preuve de discernement par rapport aux différents sites; certaines sources seront fiables alors que d'autres, plus douteuses, ne devront pas être prises en compte. Une telle méthode d'investigation permet d'établir, au-delà de la fréquence d'une expression, son utilisation plutôt locale ou étendue à l'ensemble de la francophonie (en vérifiant le pays d'origine de chaque site). Par exemple, l'expression *avenues de recherche* se retrouve 853 fois sur des sites canadiens (.ca), alors qu'une recherche sur les sites français (.fr) ne retourne que 39 occurrences (requête effectuée sur *Google* le 5 novembre 2005); nous pouvons ainsi conclure que cette expression constitue un québécoisme et qu'elle est marquée par rapport au français international standard. Nous avons eu recours à l'Internet seulement dans les cas litigieux où nos autres ouvrages de référence ne fournissaient pas de réponses adéquates.

Comme nous nous sommes servi des outils de référence pour nous aider à trancher sur le statut d'un problème lexical (voir II-2.4), nous croyions qu'il s'imposait d'utiliser plusieurs sources. Les étudiants étant déjà familiers avec *Le Petit Robert* et le

Multidictionnaire, ces ouvrages ont été retenus d'emblée. Dans le but de ne pas être trop normatif, nous avons ajouté une troisième source de référence pour bénéficier d'un avis de plus, notamment pour les québécismes, les anglicismes et les registres de langue : le logiciel *Antidote Prisme*. Nous avons jugé correct un usage considéré comme tel par l'une ou l'autre des sources de référence. Comme mentionné plus haut, les collocations apparaissant dans le *Dictionnaires des cooccurrences* de Jacques Beauchesne ont été considérées comme acceptables.

2.5 L'OUTIL D'ANALYSE : LA TYPOLOGIE DES PROBLÈMES LEXICAUX

La section suivante discute de l'outil qui nous a servi à coder les problèmes recueillis dans les textes mentionnés plus haut : la typologie des problèmes lexicaux.

2.5.1 TYPOLOGIE DESCRIPTIVE OU EXPLICATIVE?

À la section II-2.3 de notre cadre conceptuel, nous avons couvert les différents aspects du lexique, qui constituent selon nous autant de possibilités d'erreurs. Avant de proposer une typologie des problèmes lexicaux, une question importante doit être soulevée : désirons-nous une typologie exclusivement descriptive qui classe les dysfonctionnements selon le type de problème, ou alors une typologie à visée explicative qui tente – lorsque c'est possible – d'établir, au-delà du type, la source du problème observé?

Patris et Vansnick sont les seuls auteurs, à notre connaissance, à avoir proposé une typologie exhaustive des « fautes » lexicales. Ils ont élaboré leur typologie en prenant « pour base les mécanismes producteurs des dysfonctionnements observés. Que l'enseignant puisse proposer une identification de l'origine de la faute [...] semble pédagogiquement intéressant : cela devrait permettre à l'élève de prendre conscience de ce qui l'a conduit à l'erreur et dès lors de se corriger plus efficacement » (Patris et Vansnick, 1992 : 50). Nous sommes tout à fait d'accord avec l'idée d'une tentative d'explication de la source d'un problème lexical, particulièrement lorsqu'il s'agit de travailler le lexique en classe; l'analyse de problèmes lexicaux devient donc un prétexte, une amorce à la discussion. Par contre, nous croyons qu'il peut être difficile de spéculer sur la cause du dysfonctionnement alors qu'au dire même des deux

auteurs, « plusieurs hypothèses explicatives sont possibles dans bon nombre de cas » (Patris et Vansnick, 1992 : 50).

Par exemple, reprenons une erreur de régime mentionnée plus haut : **Il proteste la décision du gouvernement*. Tous s'entendront pour dire que cette phrase comporte un problème de régime puisque le caractère transitif indirect de *PROTESTER* n'est pas respecté. Par contre, il sera difficile d'identifier la source même de l'erreur. Résulte-t-elle simplement d'une connaissance approximative du verbe *PROTESTER* et de son utilisation? Provient-elle d'une contamination de l'anglais *He protests the government's decision*? Découle-t-elle plutôt d'une analogie de construction avec un quasi-synonyme : *Il conteste la décision du gouvernement*? Difficile pour un correcteur, en effet, d'identifier de façon certaine la cause d'une telle erreur.

À notre avis, aucune spéculation à propos de la source d'un problème lexical n'est pertinente avant d'avoir circonscrit de façon claire le type de problème lui-même : l'axe descriptif doit précéder l'axe explicatif. De plus, la présente recherche vise à établir un portrait général des problèmes lexicaux des étudiants universitaires et n'a pas pour but direct une intervention auprès de ceux-ci pour remédier à la situation (bien que nous souhaitions voir plus tard le développement d'une telle intervention) ; nous nous concentrons donc sur les types de problèmes plutôt que sur leurs sources²¹. Toutefois, l'idée d'analogie (phonique, sémantique, de formation, etc.) proposée par Patris et Vansnick est très intéressante et pourrait être particulièrement fructueuse pour expliquer certaines impropriétés. Nous jugeons d'ailleurs qu'une exploration dans cette direction pourrait donner lieu à une autre recherche.

2.5.2 PRÉSENTATION DE LA TYPOLOGIE ET LIMITES DE L'OUTIL

Nous sommes maintenant en mesure de proposer notre propre typologie des problèmes lexicaux²². Notre typologie se divise en trois blocs : les problèmes résultant d'une altération d'une des dimensions de l'unité lexicale, ceux relevant de l'utilisation d'une lexie fictive et ceux découlant d'un problème d'utilisation maladroite du lexique en fonction du contexte. Le premier bloc présente des regroupements entre certains types de problèmes selon qu'ils découlent d'un problème de forme, de sens, de combinatoire

²¹ Comme nous l'avons mentionné à la section II-2.3.4, les seuls types de problèmes pour lesquels nous avançons une hypothèse quant la source sont les cas d'utilisation d'une lexie fictive.

²² À la suite de notre travail d'analyse, nous avons apporté certaines modifications à notre typologie. Pour la grille révisée, voir tableau IX (IV-2.9).

grammaticale ou de combinatoire lexicale. L'ordre de ces quatre classes correspond d'ailleurs à l'organisation des informations dans un article du *Dictionnaire explicatif et combinatoire*, dont nous avons brièvement discuté dans notre cadre conceptuel (voir II-1.1.2).

Tableau III. *Typologie structurale des problèmes lexicaux*

A. Altération d'une des dimen- sions de l'unité lexicale	Dimension touchée	Transgression		Exemples
	1. Forme	Erreur d'orthographe lexicale (à l'écrit seulement).		<i>échalotte</i>
		Barbarisme : altération d'un signifiant qui se manifeste autant à l'oral qu'à l'écrit.		<i>aréoport, levier de cuisine, de la sphalte, un quipropos.</i>
	2. Sens	Non-respect d'une composante de sens	Non-respect du typage des actants.	<i>Il ne faut pas déplacer un blessé pour ne pas l'endommager.</i>
			Non-respect d'une différence spécifique	<i>Elle n'a pas entendu la sonnette, car elle somnolait profondément.</i>
	3. Propriétés de combinatoire grammaticale	Non-respect de la structure actancielle		<i>Il demande Paul de l'accompagner. Il lui a félicité pour son travail. Le bras droit a démissionné.</i>
			Utilisation d'une mauvaise préposition pour introduire les compléments. ²³	<i>Rendre visite chez des amis.</i>
		Non-respect du régime	Non-respect du mode verbal imposé par l'unité lexicale qui régit le verbe dans la subordonnée.	<i>Malgré qu'il est le seul à partir, nous devons l'appuyer.</i>
		Non-respect d'autres propriétés grammaticales inhérentes à la lexie.		<i>Une avion. La funéraille de mon oncle. Il parle lent.</i>
		Utilisation d'un mauvais verbe auxiliaire pour la formation des temps composés.		<i>J'ai parti tôt ce matin.</i>
		Non-respect du caractère figé d'une locution.		<i>Ses efforts ne valent pas la chandelle.</i>
	4. Propriétés de combinatoire lexicale	Utilisation d'un mauvais collocatif.		<i>Nous payons pour les choix que nous prenons. Heureux comme un paon.</i>

²³ Nous incluons aussi dans cette classe d'erreurs les problèmes ayant trait aux conjonctions régies par des unités lexicales. Par exemple, la phrase *Son second séjour à Londres fut une expérience meilleure *de ce qu'il avait vécu la première fois.*, l'adjectif MEILLEUR régissant la conjonction QUE.

B. Utilisa- tion d'une lexie fictive	Cas	Exemples
	1. Par la création d'une nouvelle acception d'un vocable selon un patron de polysémie productif.	<i>Il l'attendait sur la rondelle de la station Berri-UQAM</i> (= 'banc de forme arrondie')
	2. Impropropriété : association d'un sens avec le signifiant d'une lexie déjà existante, mais dont le signifié n'entretient aucun lien sémantique avec la nouvelle lexie créée.	<i>Ils ont pris un repas frugal.</i> = 'abondant, copieux'
	3. Par l'application d'une règle de dérivation morphologique généralement productive.	<i>Adapteur, parlage, moyen de transportation</i>

C. Usage douteux en fonction du contexte	Cas	Exemples
	1. Changement de registre de langue	<i>Le président a confirmé qu'il en avait plein le cul de ces rumeurs.</i>
	2. Mélange de codes (variantes dialectales)	Dans des textes en français standard : <i>Lorsqu'il faisait chaud, il sortait sa fan.</i> <i>Il l'avait rencontrée à la fin des années septante.</i>
	3. Répétition : excessive ~ en écho	<i>Le commandant commande ses troupes.</i>
	4. Imprécision : périphrase ~ mot générique	<i>Il se dirigea vers la plate-forme où les voyageurs descendaient des trains qui arrivaient.</i> = le quai d'arrivée
	5. Pléonasme	<i>Ils se réunissaient ensemble tous les jeudis.</i>

Nous croyons que la totalité des types de problèmes lexicaux possibles est couverte par cette typologie. Par contre, un tel outil présentera sans doute certaines limites lors de sa mise en application. En effet, il n'est pas toujours clair qu'un problème lexical relève d'une classe et d'une seule; par exemple, une collocation erronée peut dans certains cas être considérée aussi comme une incompatibilité sémantique découlant du non-respect du typage des actants ou d'une composante de sens. Patris et Vansnick (1992) avaient d'ailleurs émis une réserve semblable à propos de leur propre typologie. « La mise en pratique de cette typologie amène forcément l'utilisateur à prendre conscience de la difficulté de faire entrer dans des catégories bien définies une matière aussi floue et aussi fluctuante que le vocabulaire » (Patris et Vansnick, 1992 : 55). Nous sommes par contre d'accord avec eux lorsqu'ils affirment que le fait « que les catégories se combinent ou se chevauchent n'est pas un obstacle rédhibitoire dans la pratique

corrective. [...] L'essentiel est de stimuler la réflexion [sur la] faute, à partir des propositions d'identification que peut faciliter la grille catégorielle. »

2.5.3 ERREURS ET MALADRESSES : UN DEUXIÈME AXE D'ANALYSE

Cette typologie s'intéresse exclusivement à la structure même du problème et ne rend pas compte d'une autre dimension qui nous semble pourtant très importante lorsque l'on traite de lexique : la gravité du problème. Un autre axe d'analyse doit donc être pris en compte lors du codage des problèmes lexicaux. Comme cette notion est plutôt floue et qu'un jugement sur la gravité d'un dysfonctionnement est inévitablement subjectif, nous n'effectuons qu'une division approximative en deux types de problèmes lexicaux ; d'une part, les erreurs lexicales qui peuvent être condamnées catégoriquement par la norme prescriptive; d'autre part, les maladresses lexicales qui découlent d'un usage douteux du lexique, mais qui ne pourraient être de façon absolue jugées comme des erreurs. Nous désignerons ce paramètre de caractérisation du problème lexical par le terme *statut du problème*.

Nous jugeons bon introduire de telles distinctions pour rendre compte du fait que certains problèmes lexicaux sont plus « graves » que d'autres, en ce sens qu'ils contreviennent vraiment à des règles de l'usage du lexique et entrent en conflit avec l'usage standard (à l'écrit bien sûr). Les autres écarts, ceux qui provoquent un certain malaise chez le correcteur plus qu'ils ne le choquent, relèvent plutôt de la maladresse lexicale. Ces maladresses sont souvent contenues dans des tournures que nous avons peine à condamner, mais que nous ne voudrions en aucun cas voir enseignées. Comme le note Berrera-Vidal à propos de constructions syntaxiques douteuses utilisées par des apprenants d'une langue seconde, une telle erreur nous laisse perplexe, « car elle est en toute hypothèse « marquée » face au français dit standard, mais d'un point de vue strictement linguistique on ne peut pas la rejeter totalement [...] Comme on le voit, il ne s'agit plus d'une opposition fondamentale entre un emploi qui serait juste et un qui serait faux, mais d'une différence graduelle entre ce qui est toujours admis par la société et ce qui échappe parfois à la vigilance de ses censeurs [...] » (Barrera-Vidal, 1992 : 56-57).

Nous soulignons de nouveau que notre typologie n'a pas pour but premier de trancher sur le statut erroné ou non de ces maladresses lexicales, mais plutôt d'aider à diagnostiquer ce qui ne fonctionne pas dans l'utilisation discutable d'une unité

lexicale. Bien sûr, il y aura des cas où ce que nous considérons comme une erreur lexicale ne serait pas jugé comme tel par tout le monde. Par exemple, nombreux sont ceux qui seraient tentés d'accepter l'usage transitif du verbe DÉBUTER [*Il a débuté sa présentation.*], puisque de nombreuses occurrences de cette construction peuvent être relevées dans les journaux, revues et documents gouvernementaux tant au Québec qu'en France ou en Belgique, alors qu'il n'apparaît pas dans le *Petit Robert* et que le *Multidictionnaire* le condamne.

Malgré tout, nous avons quand même tenté de diminuer la part de subjectivité dans l'attribution des statuts d' « erreur » ou de « maladresse » en utilisant les ouvrages de référence. Lorsque les ouvrages indiquaient clairement une propriété de l'unité lexicale (que ce soit explicitement ou par des exemples), la transgression de cette propriété était considérée comme une erreur. Pour les cas plus flous, nous avons plutôt dû nous fier à notre jugement de locuteur, ou alors faire une recherche sur le Web ; ces cas ont la plupart du temps été considérés comme des maladreses. Notons qu'une fois de plus, l'accord interjuges a permis une certaine constance et cohérence dans le codage des problèmes relevés.

2.6 CARACTÉRISATION DU PROBLÈME EN FONCTION DU TYPE STRUCTURAL

La présente section présente en quelque sorte le mode d'emploi de la typologie structurale des problèmes lexicaux. Il s'agit en fait du pont entre les notions présentées à la section II-2.3 de notre cadre conceptuel et la méthodologie associée à l'utilisation de notre typologie. Point par point, nous caractériserons ce qui constitue un problème lexical.

A. Altération d'une des dimensions de l'unité lexicale :

1. Forme :

- i. **Orthographe lexicale** : Un mot-forme dont l'orthographe n'est pas celle des ouvrages de référence retenus entraîne une erreur d'orthographe lexicale (**hortographe*, **affiche*). Nous devons bien sûr tenir compte des nouvelles rectifications.
- ii. **Barbarisme** : Nous considérons comme un barbarisme un mot-forme dont le signifiant est déformé de telle façon que nous puissions supposer que la méconnaissance du signifiant par le scripteur se manifesterait aussi bien à l'oral qu'à l'écrit (*un *lévier de cuisine*, *un *aréoport*, **decheref*, **crocodile*).

2. Sens :

- i. **Non-respect du typage des actants** : Cette altération du sens s'observe par une incompatibilité sémantique entre le prédicat et la lexie qui remplit le rôle d'un de ses arguments. Pour les cas où le typage des actants n'est pas mentionné explicitement dans les ouvrages de référence, nous pouvons avoir recours au Web pour fin de vérification.
- ii. **Non-respect d'une différence spécifique** : C'est encore une fois par une incompatibilité sémantique que se manifeste cette transgression du sens. C'est cette fois une des différences spécifiques de la définition de la lexie qui nous permet d'en rendre compte et nous nous servons des dictionnaires pour diagnostiquer ce type de problème.

3. Propriétés de combinatoire grammaticale :

- i. **Non-respect de la structure actancielle de l'unité lexicale** (non-respect du caractère transitif (direct/indirect) ~ intransitif du verbe, omission du complément obligatoire du nom ou de l'adjectif) : Un étudiant commet une erreur relevant de cette classe 1) lorsqu'il utilise transitivement un verbe intransitif (**Le chien aboyait des jappements*), 2) lorsqu'il emploie de façon intransitive un verbe qui n'admet que l'usage transitif (**Il vexa.*), 3) lorsqu'il ne respecte pas le caractère transitif direct ou indirect d'un verbe (**Elle succéda son père au poste de directeur. *Il cherche pour un emploi.*) ou 4) lorsqu'il omet le complément obligatoire du nom ou de l'adjectif (**Le bras droit a démissionné. *Il est enclin.*). Ces informations doivent être vérifiées dans les ouvrages de référence retenus.
- ii. **Utilisation d'une mauvaise préposition pour introduire un complément** : Un problème de prépositions régies découle de l'usage d'une mauvaise préposition pour introduire un complément (*Rendre visite *chez des amis.*). Les exemples des ouvrages de référence présentent la plupart du temps les principales prépositions régies; en cas de doute, une vérification sur le Web nous permet de trancher.
- iii. **Non-respect du mode verbal imposé par la lexie qui régit le verbe de la subordonnée** : Certaines unités lexicales commandent l'emploi du subjonctif. Par conséquent, le choix de ces unités lexicales implique l'usage du subjonctif dans la subordonnée qui les suit. Il y a donc erreur lorsque ce mode verbal n'est pas respecté (**Bien qu'il est plutôt jeune... *Il recommande que vous partez plus tôt.*). Cette particularité est soulignée dans les ouvrages de référence.
- iv. **Non-respect des propriétés grammaticales inhérentes à la lexie** : Le genre nominal constitue un bon exemple de propriété inhérente à la lexie, et le non-respect d'une telle propriété entraînera une erreur lexico-grammaticale (**Une étang peu profonde.*). Chaque lexie possède, par la partie du discours à laquelle elle appartient, certaines propriétés syntaxiques qui lui permettent de se combiner d'une certaine façon avec les autres

unités lexicales. Il y a erreur lorsqu'une lexie exerce une fonction syntaxique qui n'est pas en adéquation avec sa partie du discours (**L'homme ivre parlait lent.*). Ces propriétés inhérentes à l'unité lexicale figurent toujours dans son article de dictionnaire.

- v. **Utilisation d'un mauvais auxiliaire :** Chaque verbe appelle l'usage d'un auxiliaire pour former les temps composés. L'utilisation d'un mauvais auxiliaire entraîne donc une erreur de régime (ex. **J'ai parti plus tôt.*). Les dictionnaires mentionnent généralement l'auxiliaire régi par le verbe.
- vi. **Non-respect du caractère figé d'une locution :** Les locutions sont souvent capricieuses quant à leur emploi et, les dictionnaires ne contenant à peu près pas d'informations à ce sujet, une utilisation douteuse devra être vérifiée sur le Web (*?Elle ne trouvait pas que le résultat valait la chandelle.*).

4. Propriétés de combinatoire lexicale :

Utilisation d'un mauvais collocatif : Le repérage même des problèmes de collocation comporte une certaine part de subjectivité, puisque le correcteur doit se fier à son intuition de locuteur natif pour déceler les cas problématiques. Si certaines collocations peuvent être jugées complètement erronées, parce que nettement bizarres pour un locuteur natif (ex. **effectuer une recette culinaire*), d'autres laissent le correcteur perplexe (ex. *L'homme avait de ?vastes mains.*). Nous parlerons ici aussi de maladroites lexicales plutôt que d'erreurs; il faudra juger en fonction du contexte. Nous disions plus haut que lexicque offre une certaine souplesse au locuteur, et l'emploi de collocations peu usitées peut être une façon de faire preuve de créativité (on retrouve d'ailleurs la phrase « *Il étendit à droite et à gauche ses vastes mains pour repousser les rochers croulants.* » dans *Les trois mousquetaires* d'Alexandre Dumas père!). Nous devons donc avoir recours à certains outils pour pouvoir déterminer le statut d'un problème lexical relevant de la collocation. Pour les cas litigieux, nous nous référons au *Dictionnaire des cooccurrences*; nous considérerons comme correcte une collocation y figurant. En cas de doute, nous effectuons une recherche sur le Web afin de vérifier la fréquence de la cooccurrence.

B. Utilisation d'une lexie fictive :

1. **Création d'une nouvelle acception d'un vocable :** Ce type de problème sera observé lorsque le sens de la lexie fictive utilisée sera clairement dérivé à partir de celui d'une autre lexie (par exemple par la modification de la structure actancielle, ou par une « extension » métaphorique ou métonymique du sens) et que ces deux unités lexicales partageront le même signifiant (**Il volait son cerf-volant sur la plage.*). Le dictionnaire nous permet de constater si l'utilisation de l'unité lexicale qui soulève un doute correspond à une de ses acceptions.

2. **Impropropriété** : Nous considérerons qu'il y a impropropriété lorsque la lexie utilisée présente une association signifiant ~ signifié qui ne correspond à aucune lexie du français. Par contre, le signifiant utilisé doit correspondre à une forme existante, mais associée à un ou plusieurs autres sens (*C'est un homme *intransigeant.* = 'colérique'). Encore une fois, c'est en consultant l'article de dictionnaire du vocable utilisé que nous constatons qu'il y a impropropriété.
3. **Mauvaise dérivation morphologique** : Nous sommes en présence d'une erreur de dérivation morphologique lorsqu'une forme inexistante est générée à partir d'une règle de dérivation morphologique pourtant productive en français (ex. **un adapteur, *une plante poisonneuse, *signaliser un virage*).

C. Usage douteux en fonction du contexte :

1. **Changement de registre de langue** : Les textes constituant le corpus devaient être écrits en français courant ou soutenu. Par conséquent, les lexies marquées comme familières ou vulgaires dans les ouvrages de référence sont considérées comme problématiques dans le contexte de cette recherche (*Les gens ?se foutent des lois.*).
2. **Mélange de codes** : Les anglicismes sont considérés comme des erreurs s'ils sont notés comme fautifs dans les ouvrages de référence retenus (*Il a obtenu son *degré en économie à Harvard*). Par contre, les québécoismes sont acceptés en autant qu'ils satisfassent au registre de langue imposé par le contexte.
3. **Répétition** : La répétition excessive d'une lexie ou la répétition en écho (ex. *Le commandant commande ses troupes.*) dénotent toutes deux un manque de vocabulaire et sont considérées comme des problèmes lexicaux.
4. **Imprécision** : L'emploi d'un vocabulaire pauvre et imprécis est aussi considéré comme un problème lexical. Par contre, cet aspect du lexique est souvent contrôlé de manière qualitative dans les évaluations, en se basant sur une impression générale du correcteur sur l'ensemble du texte. Dans le cadre de la présente recherche, nous notons un problème d'imprécision seulement dans les cas où une impression de lourdeur ou de flou découle directement de l'emploi d'une périphrase au lieu d'un terme précis (*Les gens se bouscullaient hors de la porte par laquelle on accède au métro.* = *bouche de métro*) ou d'un terme vague qui aurait pu être remplacé par un plus précis (*Il creusait le sol avec une chose pointue.*).
5. **Pléonasme** : Nous considérons comme un pléonasme la cooccurrence de deux unités lexicales telle que le sens de l'une est inclus dans celui de l'autre (*prévoir à l'avance, chuchoter à voix basse, première priorité*). Plutôt qu'une répétition de forme, le pléonasme constitue une « répétition sémantique ». Nous pourrions nous assurer qu'il y a pléonasme en vérifiant dans les ouvrages les définitions de deux termes le composant.

2.7 LE CODAGE ET LE CONTRE-CODAGE

Après avoir été recueilli dans les textes du corpus, chacun des problèmes lexicaux a été codé à l'aide de la typologie structurale des problèmes lexicaux et a reçu le statut d'erreur ou de maladresse.

Une fois notre codage terminé, nous avons soumis une partie des dysfonctionnements retenus à une collègue²⁴ aux fins de contre-codage. Elle a dû, à l'aide de notre typologie, coder 111 problèmes lexicaux, soit 28,4 % du total des problèmes recueillis dans notre corpus. Les 111 cas ont été choisis de façon aléatoire parmi les 391 relevés. Par contre, aucune erreur d'orthographe lexicale n'a été incluse dans ce sous-corpus, puisque cela aurait risqué, selon nous, de fausser l'accord interjuges. En effet, comme les problèmes d'orthographe lexicale sont fréquents, faciles à identifier et toujours perçus comme des erreurs (plutôt que des maladreses), leur inclusion dans le corpus de contre-codage aurait eu pour effet de gonfler l'accord interjuges. Par ce contre-codage, nous voulions plutôt valider les autres classes de notre grille, celles qui présentaient un potentiel de désaccord parce qu'elles étaient plus subjectives ou laissaient davantage place à l'interprétation.

2.8 L'ANALYSE DES DONNÉES ET LES CONCLUSIONS

Lorsque chacun des problèmes lexicaux recueillis dans le corpus de textes a été codé à l'aide de la typologie structurale des problèmes lexicaux et que le contre-codage a été effectué, nous avons pu procéder à l'analyse des données. À l'aide du logiciel *Microsoft Excel*, nous avons pu effectuer une analyse quantitative qui nous a permis de noter la fréquence de chaque type de problème lexical par rapport à l'ensemble. Nous avons aussi établi certaines tendances entre les types de problèmes relevés et les statuts (erreur ou maladresse). Finalement, les résultats recueillis lors de l'analyse des données nous ont permis d'élaborer le test qui nous a servi à répondre à notre deuxième objectif : la description des stratégies de résolution de problèmes lexicaux (voir III-3.2).

²⁴ La collègue ayant effectué notre contre-codage se nomme Ophélie Tremblay et est présentement étudiante au doctorat en didactique du français à l'Université de Montréal.

3. OBJECTIF 2 : DESCRIPTION DES STRATÉGIES DE RÉSOLUTION DE PROBLÈMES LEXICAUX

Ce second volet de la recherche correspond au deuxième objectif et vise à décrire les stratégies de résolution de problèmes mises en œuvre par les étudiants confrontés à un problème d'ordre lexical. Nous chercherons donc à comprendre quels sont les raisonnements de l'étudiant dans une telle situation. Notre choix méthodologique s'est porté sur la méthode des entretiens métagraphiques, puisqu'elle est probablement celle qui offre les meilleures chances de découvrir ce qui se passe dans la tête de nos sujets. Dans les pages qui suivent (III-3.1), nous justifierons notre choix méthodologique en présentant les entretiens métagraphiques, leur finalité, leurs limites. Les points III-3.2 à III-3.5 expliqueront quant à eux, par la présentation des aspects plus méthodologiques, la façon dont nous avons utilisé la technique des entretiens métagraphiques dans le cadre de notre recherche.

3.1 NOTRE CHOIX MÉTHODOLOGIQUE : LES ENTRETIENS MÉTAGRAPHIQUES

La présente section s'intéresse à différents aspects de la technique des entretiens métagraphiques. Par un détour du côté des entretiens d'explicitation, nous expliquerons le pourquoi d'une telle approche, en en décrivant le déroulement et le type d'informations qu'elle permet de recueillir. Avant de nous pencher plus spécifiquement sur l'entretien métagraphique, nous définirons les limites inhérentes à l'entretien d'explicitation et les écueils à éviter pour tenter d'en faire un outil efficace. Nous ferons ensuite un bref survol des différents travaux réalisés avec la technique des entretiens métagraphiques. Finalement, nous traiterons d'un type particulier d'entretiens métagraphiques auquel nous ferons appel dans le cadre de cette recherche : l'entretien en dyades.

3.1.1 L'ENTRETIEN D'EXPLICITATION

Il s'impose de parler de l'*entretien d'explicitation* avant de caractériser l'*entretien métagraphique*, puisque ce dernier constitue en fait un type d'entretien d'explicitation,

propre au domaine de la recherche en didactique. L'entretien d'explicitation est une technique qui a pour but l'élucidation d'une tâche, c'est-à-dire « établir, aussi finement que nécessaire, la succession des actions élémentaires [d'une tâche] de façon à obtenir une description « complète », et aussi à rendre intelligible la production d'une action particulière de manière à en comprendre l'inefficacité ou les erreurs, ou bien pour mettre en évidence ce qui en fait l'efficacité » (Vermersch, 1994 : 135). Le moyen pour parvenir à cette élucidation est de demander à un sujet de verbaliser sur une tâche précise, soit en cours de réalisation ou plus fréquemment *a posteriori*, afin d'avoir accès aux raisonnements, connaissances et procédures mis en œuvre.

Les objectifs visés par cette élucidation peuvent être multiples. L'objectif premier est bien entendu de permettre à l'intervieweur pratiquant ce genre d'entretien de recueillir des informations sur les processus cognitifs déployés en résolution de problèmes à différentes étapes de l'apprentissage, lui permettant par le fait même d'éclairer l'évolution des connaissances, ou plutôt leur construction. C'est généralement l'approche utilisée par les chercheurs en didactique. Les entretiens d'explicitation ont par ailleurs été récupérés, plus récemment, par les enseignants ou les formateurs qui y voient un outil didactique précieux qui « favorise la métacognition : chacun [étant] conduit à réfléchir à la fois sur l'objet [...] et sur son propre fonctionnement face à cet objet » (Cogis et Ros, 2003 : 93). Les entretiens d'explicitation sont donc aussi utilisés pour aider l'apprenant, et ultimement lui enseigner, à prendre conscience de ses propres processus cognitifs, donc à s'auto-informer. Autrement dit, cette « prise de conscience provoquée » (Vermersch, 1994 : 84) vise à conscientiser (dans le sens d' 'amener à la conscience') non seulement des connaissances qui pourraient être inconscientes chez l'apprenant, mais aussi l'articulation de ces connaissances dans la réalisation concrète d'une tâche. Ainsi, l'apprenant et le formateur sont plus à même de connaître les raisons de réussite ou d'échec de la tâche, tirant de la sorte avantage des erreurs et tâtonnements plutôt que de les condamner.

3.1.2 LE DÉROULEMENT DE L'ENTRETIEN

Un entretien d'explicitation exige la présence d'au moins deux personnes : un intervieweur et un interviewé. Nous verrons plus tard qu'il arrive d'avoir plus d'un sujet (voir III-3.1.8). Le sujet, ou interviewé, est au centre de l'entretien, qui vise à mettre au jour ses processus cognitifs par l'intermédiaire de ses verbalisations.

L'entretien porte sur une tâche particulière à propos de laquelle le sujet est appelé à verbaliser, en mettant l'accent sur les aspects procéduraux de ses actions. L'entretien pourra être réalisé durant la tâche même, ou alors *a posteriori*. Généralement l'entretien, qui vise à recueillir des données pour fin de recherche, sera filmé ou enregistré pour être ensuite retranscrit et analysé.

Pour sa part, l'intervieweur doit recentrer le sujet sur la tâche à accomplir lorsque nécessaire, en plus de lui fournir un environnement propice à la réalisation de la tâche, le plus près possible de l'environnement habituel. En effet, le propre de ce genre d'entretien est « d'observer les sujets dans des situations qui sont aussi proches que possible des situations sociales, [partant du principe] que les procédures cognitives sont en partie au moins déterminées par les conditions de production, qu'il s'agisse des tâches à réaliser ou de la qualité de l'environnement » (Jaffré et Ducard, 1996 : 94). En adoptant le point de vue « écologique » d'observer le sujet dans son environnement normal, nous augmentons les chances de voir apparaître les processus tels qu'ils sont effectivement mis en œuvre dans une situation habituelle. Une autre particularité des entretiens d'explicitation est de laisser le sujet s'exprimer dans ses propres mots, sans métalangue imposée; en plus de favoriser la verbalisation – qui pourrait se trouver compromise si le sujet devait manipuler des termes spécialisés avec lesquels il n'est pas à l'aise –, cette approche permet de vérifier quels éléments métalangagiers sont utilisés spontanément.

3.1.3 LE FRUIT DE L'ENTRETIEN D'EXPLICITATION : LES VERBALISATIONS

C'est par les verbalisations que l'entretien d'explicitation pourra atteindre son but d'élucidation. En effet, l'observation des résultats finaux (traces) et leur comparaison avec le résultat attendu (écart à la norme) permettent certes d'inférer le cheminement cognitif du sujet, mais seules les verbalisations émises éclaireront les processus effectivement mis en œuvre. Évidemment, certaines verbalisations seront plus utiles que d'autres dans la compréhension du raisonnement du sujet. L'entretien visant précisément à expliciter l'accomplissement d'une tâche, les verbalisations les plus précieuses seront celles qui portent directement sur le déroulement de l'action : les verbalisations descriptives.

Les commentaires d'ordre plutôt conceptuel – correspondant en quelque sorte aux connaissances déclaratives de la psychologie cognitive (règles, définitions, etc.) –

seront aussi intéressants, mais ne prendront vraiment toute leur valeur qu'une fois intégrés à la procédure. Il ne s'agit pas seulement de savoir quelles connaissances possède le sujet, mais surtout de quelle façon il les articule dans la réalisation d'une tâche concrète. Comme le mentionne Vermersch : « Dans le principe, le procédural se déduit des savoirs théoriques. Ces derniers fondent et justifient l'efficacité des pratiques. [Dans les faits], on peut inférer des propriétés du procédural les savoirs théoriques réellement fonctionnels » (Vermersch, 1994 : 50). C'est là le principal argument en faveur d'un type d'entretien visant la description complète des procédures pour comprendre le niveau de maîtrise des connaissances acquises et leurs rapports entre elles.

De la même façon, les commentaires véhiculant des jugements ou des opinions ne favorisent que très peu la compréhension des processus cognitifs, à moins de demander au sujet de définir les critères qui le poussent à émettre de tels jugements.

3.1.4 LES PROBLÈMES INHÉRENTS À L'ENTRETIEN D'EXPLICITATION

Bien que la technique d'entretien d'explicitation permette des découvertes intéressantes sur l'acquisition et l'articulation des connaissances, certaines difficultés lui sont inhérentes.

La première difficulté est sans doute le caractère peu naturel de la verbalisation. En effet, nous ne sommes pas habitués à commenter à voix haute nos stratégies de résolution de problèmes. De plus, bien qu'un effort soit fait pour le rendre le plus près possible de la situation réelle, les circonstances mêmes de l'entretien se révèlent peu habituelles pour le sujet : fait d'avoir à verbaliser à voix haute, relation dissymétrique intervieweur/interviewé (ou même apprenant/expert), fait d'être observé et filmé, etc. Ces facteurs pourront pousser le sujet à se cantonner dans des commentaires de l'ordre du jugement ou de l'opinion, ou encore à se limiter à décrire les circonstances plutôt que les procédures qu'il met en œuvre. Ce genre de verbalisations évite en effet au sujet de dévoiler vraiment ses processus cognitifs, lui permettant par le fait même de camoufler certaines lacunes et de ne pas s'impliquer directement. Il sera du ressort de l'intervieweur de faire en sorte que le sujet se sente à l'aise afin de favoriser le type de verbalisations qu'il espère obtenir, en l'occurrence des verbalisations descriptives relatives à l'action.

Une deuxième difficulté se situe au niveau même des connaissances. Berthoud (1982) distingue deux types d'activité en résolution de problèmes de langue : « l'activité épilinguistique » et « l'activité métalinguistique ». Les connaissances déjà conscientisées, comme les savoirs théoriques ou les savoirs procéduraux formalisés, relèvent de l'activité métalinguistique. Comme le sujet est conscient de ce type de connaissances pour avoir déjà réfléchi à leur propos, elles seront plus sujettes à faire l'objet de verbalisations. L'activité épilinguistique regroupe quant à elle « des savoirs en actes, préréfléchis, n'ayant pas fait l'objet d'une prise de conscience et que le sujet croit ne pas posséder ou, en tout cas, ne sait pas qu'il possède et donc ne sait pas les verbaliser tout seul » (Vermersch, 1994 : 86). Comme le soulignent Legros et Roy (1995), il peut exister, du fait de ces deux niveaux d'activité cognitive, un décalage entre la capacité de réflexion et la capacité de verbalisation. Autrement dit, tout ne sera pas verbalisé avec la même aisance, et certains processus seront complètement absents des commentaires du sujet, car il n'aura pas même conscience de les mettre en œuvre. De plus, « on sait également que les adultes alphabétisés, et encore davantage les experts en orthographe, « n'ont pas vraiment conscience des opérations mentales qui sous-tendent [leur] compétence orthographique » (Jaffré, 1995 : 120), donc qu'ils ne seraient pas nécessairement capables d'en témoigner verbalement » (Lefrançois et Montésinos-Gelet, 2004 : 6). Vermersch (1994) propose quand même quelques pistes à l'intervieweur pour tenter de contrer ce problème en posant des questions induisant des réponses descriptives qui ne demandent pas des connaissances de l'ordre du « déjà conceptualisé ».

Un dernier problème mérite d'être souligné : il s'agit du rappel des faits. L'entretien d'explicitation pouvant être mené *a posteriori*, des problèmes de mémoire peuvent survenir. En effet, il peut être ardu pour un sujet de se rappeler exactement son cheminement cognitif avec du recul. Il sera par conséquent très difficile de s'assurer que le procédural de la séquence verbalisée correspond effectivement à celui déployé au moment de la réalisation de la tâche. Réduire le temps entre l'action et l'entretien diminuera nécessairement l'impact de ce biais.

3.1.5 LE RÔLE DE L'INTERVIEWEUR

Le rôle de l'intervieweur, ou du chercheur-expérimentateur, est primordial dans le cadre des entretiens d'explicitation. Il ne consiste pas seulement à s'assurer que le sujet

comprenne ce qui est attendu de lui, mais aussi à analyser au fur et à mesure les verbalisations et les comportements de l'interviewé afin de l'aider à verbaliser, par des relances, des questions ou des encouragements. Il est aussi du ressort de l'intervieweur de recentrer le sujet sur sa tâche lorsque nécessaire, de lui fournir (comme nous l'avons vu plus haut) un environnement se rapprochant d'une situation habituelle et, bien sûr, de s'occuper des aspects plus techniques de l'entretien (démarrage de la caméra, enregistrement, etc.). Comme l'intervieweur doit nécessairement être conscient des buts visés par son entretien, de ce qu'il veut élucider, il sera important qu'il obtienne des verbalisations d'un niveau de description assez fin pour répondre à ces buts. Ainsi, un chercheur s'intéressant aux différents types d'informations recueillies par un étudiant lorsqu'il consulte un article de dictionnaire ne saurait se contenter d'une verbalisation du type : « *Il ne peut pas y avoir de complément, j'ai regardé dans le dictionnaire.* » Il cherchera plutôt à recueillir des verbalisations plus fines, divisant l'action « chercher dans le dictionnaire » en opérations plus simples.

Un autre aspect du rôle de l'intervieweur sera de se méfier de sa propre vision d'expert. En effet, il sera fréquent dans l'entretien d'explicitation – notamment dans le cadre des recherches en didactique où l'intervieweur est souvent un chercheur ou un enseignant – que la personne qui interviewe soit spécialiste du domaine faisant l'objet de l'entretien. Deux pièges devront particulièrement être évités. Le premier consiste à transposer sur le sujet sa propre façon de faire. Il ne faut pas oublier que l'entretien vise à élucider une procédure, et non pas à établir où se situe cette procédure par rapport à une norme, à « la bonne façon de faire ». La connaissance pointue d'un domaine et une solide expérience excluent d'emblée certains raisonnements, qui peuvent être tout à fait être présents chez le sujet « apprenti ». L'intervieweur doit donc sortir de son carcan d'expert afin de pouvoir explorer et observer des comportements qu'il n'aurait pas pu prévoir ou anticiper. Le deuxième piège à éviter, qui n'est au fond qu'une manifestation du premier, serait de ne pas questionner des descriptions qui semblent évidentes ou de considérer l'absence de verbalisations comme un gage de compréhension. En effet, comment s'assurer que le sujet est parvenu à une décision correcte par un raisonnement exact si ce raisonnement n'est pas mis au jour? Le rôle de l'expérimentateur consistant précisément à faire émerger ces raisonnements, même les actions les plus élémentaires pourront faire l'objet de questions, peu importe si leur justification semble évidente à travers des lunettes d'expert.

Il semble indispensable de faire quelques remarques à propos de la façon de formuler des relances en cours d'entretien. La fonction de ces commentaires peut être multiple : aider un sujet à verbaliser à voix haute, le pousser à approfondir certains aspects de sa description, le guider vers un type de verbalisations en particulier, ou tout simplement l'encourager à poursuivre dans le même sens. Ces relances doivent évidemment être manipulées avec prudence, afin de ne pas avoir l'effet contraire à celui escompté. Premièrement, elles doivent être exemptes de tout jugement pour ne pas donner l'impression au sujet qu'il fait fausse route et ainsi le bloquer dans ses explications. De plus, elles ne devraient pas non plus contenir d'alternative (*tu as choisi de faire ainsi pour cette raison ou pour celle-là?*) pour ne pas – comme nous l'avons vu plus haut – imposer une vision précise du problème au sujet, court-circuitant peut-être de la sorte un autre raisonnement en cours.

L'expérimentateur doit donc, en cours d'entretien, repérer les différents types de verbalisations émises par le sujet pour s'assurer qu'il obtient effectivement les informations voulues et « produire des relances qui canalisent, si nécessaire, l'élève vers le domaine de verbalisation du vécu de l'action, vers des informations centrées sur le procédural, en référence à une tâche réelle et spécifiée » (Vermersch, 1994 : 54). Selon Vermersch, les formulations en écho, qui consistent à reprendre les derniers mots du sujet pour le pousser à détailler un peu plus, sont particulièrement efficaces pour encourager une verbalisation qui va déjà dans le sens désiré, puisqu'elles véhiculent un encouragement implicite (« *si je répète ce qui vient d'être dit, c'est que je le trouve intéressant* »). Selon lui encore, les relances devraient, pour permettre à l'activité épilinguistique de transparaître aussi dans les verbalisations, induire « une réponse ponctuelle et descriptive » (Vermersch, 1994 : 86). Ainsi, les « pourquoi », situés hors de l'action puisqu'ils relèvent du domaine des connaissances, seraient moins efficaces que les « comment » pour favoriser la description de l'action.

Finalement, l'intervieweur ne doit pas non plus sous-estimer la valeur de certaines verbalisations qui, de prime abord, pourraient sembler plutôt stériles dans l'élucidation de l'action. Par exemple, les commentaires de l'ordre du jugement ou de l'opinion peuvent révéler des informations intéressantes si l'expérimentateur demande au sujet de déterminer les critères qui le poussent à les formuler. Les commentaires plutôt négatifs du type *Je n'y arrive pas* ou *je ne sais rien*, qui peuvent à première vue décourager un intervieweur, peuvent aussi constituer une amorce de verbalisation plus

révélatrice si ce dernier les récupère par des relances telles : « *Comment sais-tu que tu ne peux pas y arriver? Quelles informations te manque-t-il?* »

Nous voyons donc que le rôle de l'intervieweur est primordial. En étant attentif aux divers aspects de sa tâche, ainsi qu'aux écueils à éviter, l'intervieweur guidera le sujet vers la production de verbalisations vraiment pertinentes au regard du but d'élucidation visé.

3.1.6 L'ENTRETIEN MÉTAGRAPHIQUE

Nous avons cru bon faire une présentation assez détaillée des différentes facettes de l'entretien d'explicitation avant d'aborder l'entretien métagraphique en tant que tel. L'entretien métagraphique étant un type d'entretien d'explicitation, l'ensemble des informations présentées plus haut s'y rapportent aussi, et les techniques mentionnées lui sont applicables.

L'entretien métagraphique est le laboratoire d'observation de la démarche génétique qui « a pour but de décrire les processus de traitement et d'acquisition de l'information graphique dans toute leur complexité » (Jaffré et Ducard, 1996 : 94). Il s'agit d'un entretien d'explicitation au cours duquel un élève ou un étudiant est appelé à verbaliser à propos d'une tâche d'écriture ou de révision de texte.

Le terme « métagraphique » désigne « l'ensemble des activités mentales destinées à comprendre et à expliciter la raison d'être des traces graphiques, qu'elles correspondent à des segments phoniques, morphologiques ou lexicaux, pour s'en tenir aux activités centrées sur les mots écrits » (Jaffré, 1995 : 108). Par conséquent, un entretien métagraphique sera un entretien d'explicitation visant à élucider ces activités mentales, et un « commentaire métagraphique », une verbalisation émise par un sujet à propos d'un choix graphique.

3.1.7 L'ENTRETIEN MÉTAGRAPHIQUE COMME OUTIL DE RECHERCHE

La technique des entretiens métagraphiques a été notamment mise de l'avant par Jaffré et David durant les années 90 au sein du groupe de recherche LÉA-CNRS, qui « s'est donné comme objectif, pour partie au moins, de construire un modèle d'acquisition de l'orthographe du français » (David et Jaffré, 1997 : 155). Par conséquent, la plupart des recherches ayant utilisé cette méthode portent sur l'acquisition de la compétence

orthographique. Par exemple, Ros Dupont (1997) et Jaffré et David (1999) ont travaillé sur l'expression graphique du nombre chez les enfants de 4 à 11 ans, établissant entre autres que l'accès à ce concept morphosyntaxique est tout d'abord sémantique pour eux, autant auprès des noms que des adjectifs ou même des verbes. Lefrançois (2003) s'est pour sa part intéressée à la résolution de problèmes homophoniques chez les élèves du primaire, démontrant « que la production de graphies correctes pour des homophones est, de la 2^e à la 5^e années du primaire, surtout associée à des stratégies d'ordre métalinguistique et que le recours au métalangage ne fait une contribution utile à ces stratégies que pour certains groupes d'homophones » (Lefrançois, 2003 : 75).

Un certain nombre d'autres recherches utilisant la technique des entretiens métagraphiques se sont intéressées à différents aspects de l'acquisition de la norme orthographique chez les enfants, certaines concernant aussi la résolution de conflits sociocognitifs et la conscience morphogrammique chez les enfants de maternelle (Montésinos-Gelet et Morin : 2001).²⁵ Outre les conclusions particulières à chacun de ces travaux, les entretiens métagraphiques ont surtout permis de démontrer « que les élèves ont une activité métalinguistique indiscutable [...] ; que des obstacles existent, liés aux capacités cognitives générales ou à la complexité linguistique de l'objet ; qu'orthographier est une activité de résolution de problème et non d'application mécanique de règles » (Cogis et Ros, 2003 : 95).

Peu de recherches ayant recours à la technique des entretiens métagraphiques ont été menées auprès d'adultes, peut-être justement parce que l'acquisition et la construction de la compétence orthographique sont, après l'école secondaire, considérées comme terminées. Pourtant, les commentaires métagraphiques fournis par une population adulte permettent de constater de quelle façon s'articulent les connaissances qui ont été (ou auraient dû être) acquises tout au long des enseignements primaire et secondaire, pour éventuellement remettre en question certaines pratiques didactiques. Rappelons quand même de nouveau les deux recherches, réalisées au Québec, qui se sont intéressées à cette question.

Premièrement, la recherche menée par Roy *et al.* (1995), qui a utilisé les commentaires métagraphiques de 162 étudiants du cégep et de l'université pour dresser un bilan de leur maîtrise du code écrit. Cette étude arrive à la conclusion majeure que plus du tiers des étudiants aux ordres supérieurs d'enseignement ne possèdent pas une maîtrise suffisante du code écrit; ils ont des stratégies de résolutions de problèmes linguistiques

²⁵ Pour un survol des différentes recherches ayant utilisé la technique des entretiens métagraphiques, voir Lefrançois et Montesinos-Gelet (à paraître).

mal adaptées au contexte. En effet, les connaissances déclaratives, et dans une moindre mesure procédurales, semblent dominer dans les processus mis en œuvre devant un problème linguistique; l'absence de prise en compte et d'analyse du contexte (connaissances conditionnelles) est à la source de l'inefficacité des stratégies. De plus, la recherche révèle que la verbalisation et la réflexion sur la langue sont entravées par une maîtrise approximative de la métalangue.

La deuxième étude (Lefrançois, 2004) s'est intéressée aux stratégies de résolution de problèmes linguistiques de 24 étudiants universitaires inscrits en sciences de l'éducation à l'Université de Montréal. Cette recherche a démontré que les défaillances des stratégies déployées par les étudiants étaient dues non seulement à des problèmes au niveau des connaissances orthographiques, mais aussi à des problèmes d'accès à ces connaissances. De plus, elle a mis en lumière la difficulté des étudiants à utiliser efficacement les ouvrages de référence mis à leur disposition.

Comme nous pouvons le constater, la technique des entretiens métagraphiques a surtout été mise à profit dans l'élucidation de l'acquisition de la compétence orthographique. Nous pouvons néanmoins supposer qu'elle pourrait permettre d'intéressantes découvertes dans d'autres champs de la compétence linguistique en général : syntaxe, lexique, cohérence textuelle, etc. Une recherche utilisant les commentaires métagraphiques relatifs à ces trois aspects de la langue est d'ailleurs en cours à l'Université de Montréal, sous la direction de Pascale Lefrançois.

3.1.8 L'ENTRETIEN MÉTAGRAPHIQUE EN DYADE OU EN TRIADE

L'entretien métagraphique peut être mené avec deux ou trois sujets à la fois, par exemple en demandant à des élèves de corriger ensemble un texte composé par l'un d'eux. Ces entretiens en dyade ou en triade installent une dynamique différente et permettent de contourner certains problèmes relatifs à l'entretien en tête à tête avec l'expérimentateur.

Premièrement, travailler en groupe brise la relation dissymétrique apprenant/expert qui peut, chez certains sujets, constituer un obstacle à la verbalisation. En discutant avec des camarades de classe de son niveau, l'élève se sentira souvent plus à l'aise. Le fait que l'attention de l'intervieweur ne se porte pas sur un seul sujet peut aussi dissiper l'effet « paralysant » de l'entretien. De plus, les sujets pouvant se relancer entre eux,

les interventions de l'expérimentateur se feront nécessairement plus rares, et donc sa présence plus discrète, ce qui peut aussi contribuer à la verbalisation.

Un autre aspect intéressant des entretiens de groupe tient dans les interactions entre les sujets, qui peuvent les pousser à verbaliser davantage. En effet, il ne s'agit plus ici de simplement décrire nos propres processus, mais bien d'expliquer à l'autre notre cheminement et de le convaincre d'adopter notre solution. Dans ces entretiens, « l'interaction joue un rôle primordial : découvrant un point de vue autre dans une graphie différente, les élèves s'interrogent, s'expliquent, opposent leurs façons de faire » (Jaffré, 1997). La verbalisation devient alors plus naturelle et, conjuguée avec le fait que les élèves s'expriment hors de tout cadre métalangagier précis, peut-être plus susceptible de faire émerger l'activité épilinguistique des sujets.

Comme nous savons que le lexique suscite généralement peu de verbalisations de la part des sujets (Lafontaine et Legros, 1995 ; Roy, 1995), nous avons choisi de réaliser les entretiens métacographiques avec deux sujets à la fois afin que le fait d'avoir à confronter leurs opinions à propos de problèmes lexicaux les pousse à verbaliser davantage.

3.2 L'INSTRUMENT DE CUEILLETTE DES DONNÉES

L'instrument de cueillette de données utilisé lors des entretiens métagraphiques a été élaboré à la suite de l'atteinte de notre premier objectif de recherche : la description des types de problèmes lexicaux des étudiants universitaires. Le test est constitué de deux parties.

La première partie du test (voir test complet, annexe II) est un court texte de 20 lignes dans lequel ont été introduits 17 problèmes lexicaux. La deuxième partie du test comporte pour sa part quinze phrases n'ayant pas de liens entre elles et présentant chacune un problème lexical en caractère gras. Deux raisons expliquent que nous ayons choisi de diviser le test en deux parties. La première partie nous permettra de vérifier si les étudiants repèrent des problèmes lexicaux dans un texte et quels types sont principalement relevés. Comme nous savons par ailleurs que le lexique constitue l'une des dimensions les moins prises en compte par les étudiants lors d'une activité de révision (Roy *et al.*, 1995 ; Lefrançois, 2004), nous nous assurons, en pointant les problèmes dans la deuxième partie du test, d'avoir des verbalisations qui concernent réellement des problèmes de nature lexicale. Si nous commençons le test par le volet

de repérage d'erreurs lexicales dans un texte, c'est que nous ne voulons pas attirer l'attention des sujets sur certains types de problèmes particuliers, ce qui serait forcément arrivé si la partie « problèmes ciblés » avait précédé.

Nous voulions que les problèmes présents dans notre test respectent les proportions révélées par les conclusions de notre premier objectif. Comme nous le verrons plus loin (voir IV-1.4) les erreurs d'orthographe lexicale étaient assez fréquentes dans notre corpus de textes d'étudiants, mais étant donné que nous désirions surtout analyser les stratégies de résolution relatives aux autres types de problèmes lexicaux, c'est la seule proportion que nous n'avons pas respectée. Nous avons donc soustrait les problèmes relatifs à la forme, les cas litigieux et ceux de double codage ainsi que les erreurs d'inattention avant de rétablir les proportions des autres types de problèmes relevés dans notre corpus. Nous avons tout de même introduit une erreur d'orthographe lexicale et un barbarisme dans le test. Voici donc la répartition des 32 problèmes lexicaux introduits dans notre test : 10 problèmes de combinatoire grammaticale, 8 problèmes de mauvaise utilisation d'unité lexicale en fonction du contexte, 6 cas d'utilisation d'une lexie fictive, 3 problèmes de combinatoire lexicale et 3 problèmes de sens. L'ensemble des problèmes a été tiré directement des textes de notre corpus d'analyse du premier objectif ou s'en sont fortement inspirées. Comme nous ne voulions pas introduire de cas trop litigieux, nous avons tenté de retenir surtout des erreurs, plutôt que des maladresses, pour construire le test; les quelques cas pour lesquels nous n'avions que des maladresses ont été concentrés dans la deuxième partie du test pour que les sujets n'aient pas à les repérer eux-mêmes.

3.3 LA CONSTITUTION DU CORPUS DE VERBALISATIONS

C'est à partir d'entretiens métagraphiques menés auprès d'étudiants universitaires que nous avons amassé les données qui nous ont servi à répondre à notre deuxième objectif, soit la description des stratégies de résolution de problèmes lexicaux déployées par ces étudiants.

3.3.1 L'ÉCHANTILLON

L'échantillon est composé de 16 sujets qui ont travaillé par deux; il comporte 11 femmes et 5 hommes. Tous les sujets étaient inscrits au baccalauréat au moment de la

passation du test et proviennent de différents programmes : sciences biomédicales, physiothérapie, pharmacie, communication, sécurité et études policières, architecture, études françaises, anthropologie, administration, adaptation scolaire, science politique et histoire de l'art. Le recrutement des sujets a été compliqué par la grève des étudiants au mois d'avril 2005. Par conséquent, les sujets ont participé à l'étude sur une base volontaire et aucune information relative à leur compétence en français écrit n'a pu être obtenue²⁶. Malgré tout, nous pouvons évaluer, après avoir vu les étudiants réviser les textes et réfléchir sur des problèmes linguistiques, que cet échantillon possède une maîtrise satisfaisante du code écrit (aucun sujet ne nous a semblé très faible).

3.3.2 L'ADMINISTRATION DES TESTS

Les entretiens métagraphiques ont été réalisés à la Faculté des sciences de l'éducation par l'auteur de la présente recherche²⁷. Tous les entretiens ont été filmés et enregistrés sur cassette audio. En plus de nous assurer une copie de rechange en cas de problème avec la cassette audio, l'enregistrement vidéo des entrevues nous a permis d'observer la façon dont les sujets utilisent les ouvrages de référence.

Pour la première partie du test, les étudiants devaient lire le texte et discuter des problèmes qu'il présentait selon eux. Bien sûr, le nombre de problèmes présents dans le texte n'a pas été mentionné, ni le fait que l'ensemble soient d'ordre lexical, afin de ne pas guider l'attention des sujets vers des problèmes particuliers. Ces derniers ont aussi été invités à proposer une ou plusieurs solutions pour rendre correct chacun des problèmes relevés. Des ouvrages de référence (*Petit Robert* et *Multidictionnaire*) étaient à la disposition des sujets.

La deuxième section du test portant sur des problèmes ciblés, les étudiants n'avaient qu'à discuter du problème mentionné pour tenter de l'expliquer et à proposer une correction, encore une fois en ayant la possibilité de recourir aux ouvrages de référence.

Le fait de travailler en dyade a poussé les étudiants à verbaliser (voir III-3.1.8), mais l'intervieweur a parfois dû intervenir et relancer les sujets lorsque ceux-ci semblaient dans une impasse. Les étudiants ont pu bénéficier du temps désiré pour accomplir les

²⁶ Comme pour le premier objectif, les 16 sujets qui ont participé aux entretiens métagraphiques ont rempli des formulaires de consentement. Une copie de ce formulaire figure dans l'annexe I.

²⁷ Pour les consignes exactes données aux étudiants sur la tâche à accomplir, voir annexe III.

deux tâches, mais l'ensemble des équipes ont complété le test en moins d'une heure quinze minutes.

3.3.3 LA TRANSCRIPTION DES ENTREVUES

Les huit entretiens métagraphiques réalisés ont été transcrits intégralement. Chacune des interventions faite par les étudiants occupe une ligne différente et est associée au sujet qui l'a émise (même si cette variable n'a finalement pas été prise en compte lors du codage). Les interventions de l'intervieweur ont été notées en italique. Les recours aux ouvrages de référence et les hésitations ont aussi été relevés.

3.4 LE CODAGE DES DONNÉES

Aux fins de codage, les entretiens retranscrits ont été découpés en séquences. Une séquence est un ensemble de commentaires se rapportant à un problème particulier. Comme nous ne désirons pas faire des portraits individuels des sujets, nous avons regroupé dans une même séquence les commentaires des deux membres de l'équipe. Notre codage comporte donc deux niveaux, auxquels correspondent deux grilles différentes : le codage des séquences et le codage des commentaires.

3.4.1 LE CODAGE DES SÉQUENCES

Voici les critères retenus pour l'analyse des séquences :

Tableau IV. Grille de codage des séquences

Rubriques	Catégories
1. séquence traite d'un problème cible	<i>oui</i>
	<i>non</i>
2. type de problème traité	<i>problème de forme</i>
	<i>problème de sens</i>
	<i>problème de combinatoire grammaticale</i>
	<i>problème de combinatoire lexicale</i>
	<i>utilisation douteuse d'une unité lexicale en fonction du contexte</i>
	<i>lexical non défini</i>
	<i>relève d'une dimension non lexicale</i>
3. statut du problème au départ	<i>erroné</i>
	<i>non erroné</i>
4. statut du problème à la fin	<i>erroné</i>
	<i>non erroné</i>
5. changement	<i>effectué</i>
	<i>non effectué</i>
6. contenu de la séquence	<i>suffisant</i>
	<i>insuffisant</i>

Le premier critère (*problème cible*) n'a servi à coder que les séquences de la première partie du test. Il spécifie si la séquence porte effectivement sur l'un des problèmes lexicaux introduits dans le texte, ou alors sur un autre élément. Ce critère n'est pas pertinent pour le codage des séquences issues de la deuxième partie du test, puisque les problèmes y figurant sont en caractères gras; les échanges ne traitent par conséquent que des problèmes cibles.

Le deuxième critère (*type de problème traité*) nous permet de constater quelles dimensions lexicales suscitent les commentaires chez les sujets. Lors de l'analyse, il nous a aussi aidé à voir les types de stratégies de résolution mises en œuvre pour différents types de problèmes traités. Comme le nombre de problèmes lexicaux présents dans le test est assez restreint (32), nous n'avons pas cru nécessaire d'accorder autant de valeurs possibles au critère 2 qu'il en existe dans notre typologie structurale des problèmes lexicaux. Nous avons plutôt choisi de conserver les cinq grandes divisions de la typologie (problèmes de forme, de sens, de combinatoire grammaticale, de combinatoire lexicale, d'utilisation douteuse d'une lexie en fonction du contexte)

afin d'alléger notre codage. Par contre, nous avons en cours d'analyse introduit des distinctions plus fines pour certaines de ces grandes classes d'erreurs lexicales afin d'observer de quelle façon des problèmes particuliers, par exemple des pléonasmes, étaient analysés et corrigés.

Comme il était évident qu'aucun des cas 6 d'utilisation d'une lexie fictive ne serait abordé comme tel par les sujets, nous les avons répartis dans les classes de problèmes lexicaux qui pouvaient le mieux leur être associées. Ainsi, les 4 impropriétés et le cas de création d'une nouvelle acception d'un vocable présents dans le test ont été associés aux 3 problèmes de sens déjà présents pour l'analyse (ce qui porte le nombre de problèmes de sens à 8). Le cas de mauvaise dérivation morphologique inséré dans le test est considéré comme un problème de forme dans l'analyse. Le barbarisme que nous avons retenu pour le test, *tant qu'à moi*, a été perçu par beaucoup de sujets comme une unité lexicale appartenant au registre familier. Par conséquent, nous l'avons considéré comme un problème de registre de langue lors du codage des entrevues. Il est aussi important de mentionner que les problèmes cibles ont toujours été codés par la classe d'erreur qui correspondait à notre analyse de l'objectif 1, et ce peu importe la perspective d'analyse adoptée par les sujets. Par exemple, si un problème de combinatoire lexicale (collocation) était traité comme un anglicisme par les sujets, nous aurions tout de même codé la séquence correspondante *combinatoire lexicale*.

Comme certaines séquences portent sur des problèmes lexicaux non ciblés et qu'il n'est pas toujours facile de définir à quelle dimension de l'unité lexicale elles font référence, nous avons ajouté un code *lexical non défini*. Nous avons utilisé ce code toutes les fois où il était impossible de caractériser plus en détail la séquence; dans les cas où il était clair que la séquence s'intéressait à une dimension particulière, nous lui avons attribué le type de problème correspondant.

Par ailleurs, la présente recherche s'intéressant aux stratégies de résolution de problèmes lexicaux exclusivement, et le test ne présentant que des problèmes de ce type, toutes les séquences codées *relevant d'une dimension non lexicale* n'ont pas été retenues pour l'analyse.

Les critères 3 et 4 (*statut du problème au début/à la fin*) fonctionnent en paire et permettent de constater l'évolution du produit (ce qui se retrouve sur la feuille) : le problème considéré était-il effectivement erroné au départ? Est-ce que, à l'issue de la séquence, le statut demeure le même? Si une modification a été apportée, a-t-elle

permis de rectifier le problème? Pour les quelques cas où la solution correcte présentait une erreur de nature non lexicale (mauvais accord), nous avons tout de même considéré que le problème avait été résolu.

Le cinquième critère (*changement effectué ~ non effectué*) peut à première vue sembler redondant par rapport aux deux critères précédents. En effet, si un problème est *erroné* au départ et *non erroné* à la fin de la séquence, il est évident qu'un changement a été apporté. La même inférence est valable dans le cas où le problème est *non erroné* au départ et *erroné* à la fin de la séquence. Par contre, les seules valeurs présentes dans les critères 3 et 4 ne permettent pas de rendre compte de cas où une erreur (*problème erroné*) est, à l'issue de la séquence, remplacée par une autre erreur. De la même façon, une séquence peut porter sur un problème *non erroné* et déboucher sur un changement par un autre élément, correct lui aussi, qui laisse le statut du problème intact. C'est pour cette raison que le critère 5 a aussi été introduit dans la grille d'analyse.

Finalement, **le critère 6** (*contenu de la séquence*) précise si les stratégies contenues dans la séquence ont été suffisantes pour résoudre correctement le problème. Nous avons considéré qu'une séquence présentait des stratégies suffisantes lorsque l'explication fournie par les sujets aurait pu être acceptable pour expliquer le problème ciblé à quelqu'un ayant le français comme langue seconde ou à un apprenant francophone confronté à ce problème. Il n'était donc pas nécessaire qu'une séquence contienne des termes propres au métalangage lexical pour être codée *suffisante*; il fallait simplement qu'elle dénote une bonne intuition dans l'analyse du problème traité. Dans certains cas, par exemple les problèmes de forme ou de sens, nous avons considéré suffisant le simple fait de préciser l'orthographe correcte ou d'énoncer le sens d'une unité lexicale.

3.4.2 LE CODAGE DES COMMENTAIRES

Bien entendu, tous les commentaires émis par les sujets n'ont pas été codés. Premièrement, nous avons éliminé tous les commentaires appartenant aux séquences non lexicales et ceux qui n'étaient pas reliés à la tâche. Les commentaires incomplets ou inintelligibles ainsi que ceux témoignant des interactions entre les sujets n'ont pas été codés, mais ont tout de même été conservés afin que les commentaires retenus pour le codage apparaissent dans le contexte dans lequel ils ont été émis. Finalement, les

commentaires dont le contenu se répétait à l'intérieur d'une même séquence n'ont été codés qu'une seule fois par séquence.

Les critères retenus pour l'analyse des commentaires figurent dans la grille présentée plus bas (tableau V).

Le premier critère renseigne sur la *spontanéité du commentaire*. Ont été considérés comme spontanés²⁸ les commentaires émis de façon naturelle par les sujets, ou ceux issus d'une interaction entre ces derniers. Les commentaires suscités sont ceux formulés en réaction à une intervention de l'intervieweur. Nous avons distingué deux types de commentaires suscités. Les premiers sont les « commentaires d'approfondissement », c'est-à-dire ceux émis en réaction à une intervention de l'intervieweur demandant à expliciter un choix ou une décision, par exemple *Pourquoi as-tu changé ce mot?* Les seconds commentaires suscités sont ceux faits à la suite d'une question de l'intervieweur contenant clairement une piste d'analyse ou un élément de réponse, par exemple *Crois-tu que si tu cherchais plutôt le mot X dans le dictionnaire, tu risquerais davantage de trouver la réponse?* Ont aussi été considérés « suscité – piste d'analyse » les commentaires dans lesquels les sujets évaluaient une solution proposée par l'intervieweur, comme *Est-ce que ça aurait été correct si on avait plutôt dit X?* Nous croyons pertinent d'introduire cette distinction, car il est fort probable que les interventions de l'intervieweur aient guidé les sujets vers des réflexions ou des pistes d'informations auxquelles ils n'auraient possiblement pas eu accès sans ces interventions. Ainsi, les commentaires suscités témoignent peut-être moins de la réaction habituelle d'un étudiant face à un problème lexical, mais plutôt de stratégies de résolution auxquelles l'étudiant n'aurait pas pensé de lui-même.

²⁸ Bien que cette distinction entre les commentaires suscités et les commentaires spontanés soit reconnue et considérée par la plupart des chercheurs, nous pouvons observer certaines divergences terminologiques. Par exemple, Morin (2002) parle de *commentaires naturels* plutôt que de *commentaires spontanés*. Nous croyons pour notre part que le terme *spontané* est plus parlant et c'est celui que nous retiendrons.

Tableau V. Grille de codage des commentaires

Rubriques		Catégories
1. spontanéité du commentaire		<i>spontané</i>
		<i>suscité – approfondissement</i>
		<i>suscité – piste d'analyse</i>
2. statut du commentaire		<i>approprié</i>
		<i>non approprié</i>
		<i>indéfini</i>
3. étape de la résolution du problème		<i>repérage du problème</i>
		<i>analyse du problème</i>
		<i>proposition et analyse de solutions</i>
		<i>prise de décision</i>
4. présence de métalangage		<i>oui</i>
		<i>non</i>
5. types de commentaires	commentaires visant la résolution du problème	<i>proposition catégorique (non justifiée)</i>
		<i>proposition faisant partie d'un raisonnement (justifiée)</i>
		<i>référence à la forme du mot-forme</i>
		<i>recours au sens</i>
		<i>établissement d'un lien morphologique avec d'autres mots-formes</i>
		<i>référence à des propriétés de combinatoire grammaticale</i>
		<i>recours au jugement du locuteur natif</i>
		<i>référence à une autorité</i>
		<i>recours à la paraphrase, la reformulation</i>
		<i>remplacement par une autre unité lexicale</i>
		<i>comparaison de deux solutions (oral ou écrit)</i>
		<i>consultation d'un ouvrage de référence</i>
		<i>analyse grammaticale ou sémantique de la phrase</i>
		<i>référence à la norme fonctionnelle</i>
		<i>comparaison/analogie avec un autre problème</i>
		<i>établissement d'un lien avec une unité lexicale de l'anglais</i>
		<i>référence à une expression</i>
		<i>énonciation d'une « règle »</i>
		<i>apport d'une évidence linguistique</i>
		<i>illustration d'une bonne utilisation de l'unité lexicale traitée par un exemple</i>
		<i>mise en parallèle d'un bon et d'un mauvais exemple</i>
		<i>relecture</i>
		<i>évitement</i>
		<i>autre</i>
	commentaires diagnostiques	<i>jugement global</i>
		<i>jugement esthétique</i>
	commentaires périphériques	<i>commentaire relatif à la tâche</i>
		<i>expression d'incertitude</i>

Le deuxième critère (*approprié* ~ *non approprié*) renvoie au contenu même du commentaire. Un commentaire approprié véhicule un contenu 1) qui est en adéquation avec le système de la langue et 2) qui s'avère pertinent dans la résolution du problème

traité. Inversement, nous avons considéré comme non appropriée une remarque 1) dont le contenu contredit la norme lexicale telle que nous l'avons définie ou 2) qui n'est d'aucune aide dans la résolution du problème qui en fait l'objet. Ce paramètre de codage correspond au critère *normé ~ non normé* utilisé par Lefrançois (2004) dans sa grille d'analyse. La raison pour laquelle nous préférons ici le terme *approprié* à celui de *normé* est que notre recherche s'intéresse au lexique en particulier et que nous avons vu la difficulté d'établir une norme lexicale. Nous croyons que le terme *approprié* traduit mieux les limites floues de certains aspects de cette norme. Comme il était impossible de statuer sur le caractère approprié ou non de certains commentaires – notamment les questions – nous avons aussi introduit une valeur *indéfini*.

Le troisième critère indique à quelle étape de la résolution du problème correspond le commentaire. Les valeurs possibles correspondent aux quatre étapes définies dans notre cadre conceptuel (voir II-3.3.3). Dans les cas où plusieurs étapes semblaient se chevaucher dans un même commentaire, nous avons associé ce dernier à l'étape la plus avancée du processus de résolution. Par exemple, un commentaire relevant à la fois du repérage et de l'analyse du problème a été codé *analyse du problème*. L'analyse d'une solution proposée pour régler le problème discuté appartient à la phase de *générat[i]on, applicat[i]on et analyse d'une solut[i]on* et non à celle d'analyse.

Le quatrième critère mentionne si le commentaire comporte des éléments de métalangage. Lors de l'analyse, il nous permettra de déterminer quels sont les éléments de métalangage utilisés par les étudiants. Il est à noter que les termes de métalangage très fréquents dans le langage courant (MOT, PHRASE, SENS) n'ont pas été considérés comme des recours au métalangage.

Le deuxième objectif de la présente recherche vise la description des stratégies de résolution de problèmes lexicaux d'étudiants universitaires. Par conséquent, **le critère 5 (type de commentaires)** constitue le cœur de notre grille d'analyse, puisque c'est précisément celui qui nous a permis de voir de quoi sont constituées les stratégies mises en œuvre. Si nos deux grilles de codage sont plutôt de type fermé parce qu'elles contiennent « des rubriques, des catégories et des valeurs dont la liste est définie dès la construction du cadre conceptuel et méthodologique » (Van der Maren, 1996 : 436), le critère 5 constitue la seule rubrique ouverte au départ, c'est-à-dire que nous avons été amené à y ajouter certaines valeurs durant le codage. Comme la recherche vise à définir les types de stratégies de résolution déployées, il est tout à fait logique que

notre liste n'ait pas été exhaustive dès le début de notre analyse. Voici un bref survol des types de commentaires apparaissant dans notre grille de codage :

- *Proposition catégorique* : cette stratégie est parmi les moins sophistiquées. Le sujet propose une solution au problème observé sans aucune analyse ni justification ;
- *Proposition faisant partie d'un raisonnement* : un tel commentaire contient une proposition de solution au problème, mais qui s'inscrit dans un raisonnement sur le problème traité. Comme les sujets avaient pour consigne de proposer des solutions aux problèmes discutés, les commentaires recevant ce code devraient être assez fréquents et ne seront pas d'un grand intérêt lors de l'analyse ;
- *Référence à la forme du mot-forme* : ce type de verbalisation porte sur le signifiant d'un mot-forme et consiste généralement à en spécifier l'orthographe correcte (ou présumée correcte) ;
- *Recours au sens* : cette stratégie consiste à faire référence au sens d'une unité lexicale, mais sans avoir recours au dictionnaire. En d'autres mots, il s'agit de donner une définition personnelle d'un terme, ou de faire référence à une composante de sens en particulier ;
- *Établissement d'un lien morphologique avec d'autres mots-formes* : cette stratégie consiste à établir un lien entre deux mots-formes en ce qui concerne leurs signifiants (souvent dans le but de vérifier l'orthographe de l'une d'elles) ;
- *Référence à des propriétés de combinatoire grammaticale* : le commentaire recevant ce code contient une référence (explicite ou implicite) à une propriété de combinatoire grammaticale. Par exemple, il mentionne que le verbe ciblé est transitif, qu'un nom est invariable, etc. ;
- *Recours au jugement du locuteur natif* : un tel commentaire contient une appréciation à propos d'une formulation basée sur une simple intuition ;
- *Référence à une autorité* : un cas typique de cette stratégie serait, par exemple, de mentionner qu'un professeur condamne tel emploi d'une unité lexicale sans expliquer pourquoi (*on m'a déjà dit de pas utiliser ce mot-là*). Par contre, nous avons aussi considéré comme des références à une autorité des commentaires comme *J'ai entendu à la radio que...* ou *J'ai déjà vu à l'école qu'on n'avait pas le droit de...* ;
- *Recours à la paraphrase, à la reformulation* : cette stratégie consiste à paraphraser ou à reformuler une phrase dans le but de mieux en comprendre le sens ou la structure ;
- *Remplacement par un autre terme* : le sujet remplace le terme ciblé par une autre unité lexicale afin de pouvoir mieux juger de sa pertinence ;
- *Comparaison de deux solutions* : cette stratégie consiste à mettre en opposition deux solutions possibles (soit à l'oral ou par écrit) pour juger de celle à privilégier ou pour mettre en évidence des relations d'antonymie ou de synonymie ;
- *Consultation d'un ouvrage de référence* : ce code sert à identifier les commentaires présentant la lecture d'un article (ou d'une partie d'article) de dictionnaire ;
- *Analyse grammaticale ou sémantique de la phrase* : les commentaires ainsi codés portent sur une analyse de la phrase contenant le problème ciblé. Ces commentaires permettent de clarifier la structure syntaxique de la phrase ou de mettre en lumière une ou plusieurs interprétation(s) possible(s) ;

- *Référence à la norme fonctionnelle* : cette stratégie consiste à mentionner que l'emploi d'un terme particulier n'est pas acceptable dans un contexte particulier en vertu de la norme fonctionnelle. De tels commentaires portent presque exclusivement sur les problèmes de registres de langue ;
- *Comparaison/analogie avec un autre problème* : stratégie qui consiste à comparer le problème ciblé avec un autre problème déjà discuté. Par exemple, *Ça ressemble au problème de tantôt* ou *C'est comme avec « avion », qu'on sait jamais si c'est féminin ou masculin* ;
- *Établissement d'un lien avec une unité lexicale de l'anglais* : utilisée exclusivement dans les cas d'anglicismes (réels ou supposés), cette stratégie consiste à nommer une unité lexicale de l'anglais qui nous pousse à penser que l'unité lexicale traitée est un anglicisme ;
- *Référence à une expression* : un tel commentaire mentionne que l'expression utilisée dans le texte est une locution du français ou fait référence à sa forme ;
- *Énonciation d'une « règle »* : le sujet mentionne, sous forme d'une règle réelle ou inventée, une particularité de l'unité lexicale problématique ;
- *Apport d'une évidence linguistique* : cette stratégie consiste à apporter une évidence linguistique (titre de livre, parole de chanson, parole célèbre, etc.) qui prouve une certaine utilisation de l'unité lexicale sur laquelle porte le problème ;
- *Illustration d'une bonne utilisation de l'unité lexicale traitée par un exemple* : le sujet fournit un exemple d'utilisation de l'unité lexicale traitée ;
- *Mise en parallèle d'un bon et d'un mauvais exemple* : en plus de fournir un exemple correct d'utilisation d'une unité lexicale, le sujet le contraste avec un exemple qu'il juge incorrect. Ces commentaires prennent souvent la forme « *Tu peux dire X, mais tu ne peux pas dire Y* » ;
- *Relecture* consiste à relire une phrase ou un bout de phrase avec les modifications apportées ;
- *Évitement* : la stratégie d'évitement présente une reformulation complète de la phrase contenant le problème ciblé, de sorte qu'il est complètement évacué ; plutôt que de proposer une solution, le sujet fait disparaître le cas problématique.

Finalement, une classe *autre* nous a permis de regrouper l'ensemble des stratégies que nous n'avions pas prévues au départ. L'analyse des commentaires appartenant à cette classe a permis de révéler de nouvelles stratégies. La grille présentée plus haut contient d'ailleurs quelques stratégies qui ont été ajoutées en cours d'analyse. Les quelques commentaires qui demeureront dans la classe *autre* pour l'analyse sont ceux ne correspondant pas aux types de stratégies apparaissant dans notre grille de codage. Les types de commentaires que nous venons de détailler visent directement la résolution du problème et se rapprochent en ce sens de stratégies véritables. Par contre, d'autres types de commentaires figurent aussi dans les séquences et font aussi partie du processus de résolution de problème même s'ils ne contribuent pas directement à trouver une solution. Nous avons regroupé ces autres types de commentaires dans deux classes : les commentaires diagnostiques et les commentaires périphériques.

Les commentaires diagnostiques sont ceux qui véhiculent une appréciation d'un problème traité ou d'une solution proposée. Nous avons distingué deux types de commentaires diagnostiques : les jugements globaux et les jugements esthétiques. Un *jugement global* est un commentaire exprimant une appréciation catégorique d'un problème, par exemple *Ça s'est dit pas* ou *C'est correct*. Le jugement esthétique est du même ordre, sauf qu'il véhicule une appréciation d'ordre esthétique comme *C'est pas beau comme ça* ou *Ça sonne mieux*. Sans contribuer directement à la découverte d'une solution, ces verbalisations orientent le processus de résolution du problème en statuant sur la pertinence d'une solution ou le statut erroné ou non d'un élément spécifique. Par exemple, ce sera souvent un commentaire diagnostique qui marquera la fin d'une séquence ou qui indiquera que la solution proposée est acceptée ou que la recherche de meilleures solutions doit continuer.

Les commentaires que nous qualifions de « périphériques » ne contribuent pas vraiment à la résolution du problème. Les commentaires relatifs à la tâche sont des commentaires généraux faits à propos de la difficulté du problème traité ou de la tâche à accomplir, par exemple *Ça, j'ai toujours eu de la misère avec ça* ou *Ben là, c'est facile ici*. Les commentaires ayant reçu le code *expression d'incertitude* contiennent des éléments laissant sous-entendre que le sujet n'est pas convaincu de sa position par rapport au problème.

Lors du codage des entrevues, nous avons tenté de n'attribuer qu'un seul code par verbalisation. Les seuls cas pour lesquels cette règle n'a pas été respectée sont les commentaires diagnostiques et les commentaires périphériques. La raison est simple : il arrive fréquemment que les verbalisations visant la résolution du problème véhiculent aussi ce type de commentaires. Par exemple, une proposition de solution survenant en cours d'analyse du problème pourrait être accompagnée d'une expression d'incertitude. Ainsi, nous avons dans certains cas attribué deux « types de commentaires » à une seule verbalisation.

3.5 L'ANALYSE DES DONNÉES

Une fois les entrevues codées, nous avons pu procéder à l'analyse et à l'interprétation des données. Comme nous avons codé différemment les commentaires et les séquences, nous les avons analysés séparément.

3.5.1 L'ANALYSE DES SÉQUENCES

L'analyse des séquences nous a permis de répondre à différentes questions. L'analyse de celles de la première partie du test (repérage d'erreurs) nous a d'abord permis de vérifier quels types de problèmes ont principalement été traités par les étudiants. Les étudiants ont-ils repéré les problèmes lexicaux inclus dans le texte où se sont-ils plutôt attardés à d'autres dimensions? Quels problèmes lexicaux ont été davantage repérés?

Dans un deuxième temps, l'analyse de l'ensemble des séquences du test nous a permis de constater le nombre de problème bien résolus parmi les problèmes traités. De plus, nous avons pu observer si les problèmes bien résolus correspondaient à des séquences véhiculant des stratégies de résolution de problèmes suffisantes.

3.5.2 L'ANALYSE DES COMMENTAIRES

L'analyse des commentaires nous a permis d'aller plus en profondeur dans l'observation du contenu des verbalisations recueillies lors des entrevues. Cette étape a mis en lumière les stratégies déployées par les étudiants dans la résolution de problèmes lexicaux. Par l'analyse des données, nous avons pu voir quel pourcentage des verbalisations étaient appropriées, quelles étapes de la résolution avaient occasionné le plus de commentaires et quels sont les éléments du métalangage utilisés spontanément par les étudiants dans une tâche de résolution de problème lexical. Après ces observations plus générales, nous avons regardé plus en détail le contenu des verbalisations pour déterminer les types de commentaires émis par les sujets. Finalement, nous avons isolé les stratégies de résolution dominantes pour chaque type de problème traité afin de constater si elles étaient appropriées et efficaces.

3.6 LIMITES DE LA RECHERCHE

Évidemment, tout travail de recherche présente certaines limites, et ceci est particulièrement vrai dans le cadre d'un mémoire de maîtrise pour lequel le temps et les moyens sont eux-mêmes assez limités. Il est donc important de bien délimiter notre champ d'investigation et de reconnaître les limites découlant des choix effectués. Voici donc les principales limites de la présente recherche :

- L'ensemble des textes analysés dans le cadre de notre premier objectif ont été produits par des étudiants appartenant tous à une même faculté (Faculté des sciences de l'éducation). Bien que nous soyons à peu près convaincu que notre échantillon est tout de même représentatif des étudiants universitaires en général, il y aurait lieu de s'en assurer en analysant des textes produits par des étudiants originaires d'autres facultés et départements.
- Nous avons éliminé de notre échantillon tous les étudiants n'ayant pas été scolarisés en français tout au long de leurs études. La raison pour laquelle nous avons procédé ainsi est claire : les problèmes lexicaux des étudiants allophones seraient probablement différentes de ceux des francophones, du moins en ce qui regarde leur fréquence. Refaire la même recherche avec des étudiants allophones permettrait certainement plusieurs observations très intéressantes.
- Nous prétendons implicitement évaluer la maîtrise lexicale de nos sujets, mais il faut bien entendu relativiser ces propos. En effet, nous avons vu à la section I-3.1 que l'analyse de courts textes ne permet pas de rendre compte de la maîtrise lexicale d'un individu. En fait, cette démarche ne permet d'évaluer que sa maîtrise d'une sous-partie de son vocabulaire : celui qu'il aura choisi d'utiliser dans le texte analysé. De plus, il est impossible d'affirmer, par la seule utilisation correcte d'une lexie en contexte, que le sujet observé maîtrise totalement l'ensemble des propriétés de l'unité lexicale utilisée. Par contre, certains types d'erreurs récurrentes dans le choix « arbitraire » des unités lexicales qu'opère le scripteur suggéreront que ce dernier a certaines lacunes en ce qui concerne les dimensions affectées.
- Le nombre de sujets de notre deuxième objectif (16) est très limité et ne permet de faire aucune généralisation. Compte tenu du fait que les sujets ont participé à l'étude de façon volontaire et que l'auteur de la présente recherche a dû transcrire et analyser lui-même l'ensemble des entrevues, nous croyons tout de même que l'observation de huit dyades nous a permis de dresser un portrait assez juste des stratégies de résolution déployées. Il pourrait être intéressant de refaire une étude semblable avec plus d'étudiants, peut-être en tenant compte, cette fois, de leur niveau en français écrit; peut-être serait-il ainsi possible d'observer certaines tendances entre les stratégies mises en œuvre et le degré de maîtrise du code écrit.

Si notre recherche présente certaines limites au niveau des échantillons, elle a par contre le mérite de s'intéresser à une question sur laquelle ne portent que très peu de travaux. Nous espérons donc que, malgré les limites de la zone explorée, elle aura permis de défricher une certaine partie de la question lexicale et d'ainsi ouvrir la voie à des travaux de plus grande envergure.

IV. LES PROBLÈMES LEXICAUX DES ÉTUDIANTS UNIVERSITAIRES

1. PRÉSENTATION DES RÉSULTATS (OBJECTIF 1)

Le présent chapitre vise la présentation et l'interprétation des résultats obtenus à la suite de l'analyse de 103 textes d'étudiants universitaires. Rappelons tout d'abord l'objectif visé par le premier volet de notre recherche :

Décrire les types de problèmes lexicaux d'étudiants universitaires soumis à une activité de rédaction de texte et établir la fréquence de chaque type.

Pour témoigner du fait que l'analyse de nos données a été effectuée en plusieurs étapes, la section « présentation des résultats » sera divisée en quatre parties. La première partie (IV-1.1) s'intéressera au codage des problèmes lexicaux. La deuxième partie (IV-1.2) sera consacrée à l'analyse des erreurs lexicales, alors que la troisième (IV-1.3) portera sur les maladroresses. Finalement, la quatrième partie (IV-1.4) brossera un tableau complet des différents dysfonctionnements lexicaux observés.

1.1 LE CODAGE ET LE CONTRE-CODAGE : RÉSULTATS

La correction des 103 copies d'étudiants a révélé un total de 391 problèmes relatifs au lexique, ce qui constitue une moyenne de 3,8 problèmes par copie. De ces 391 dysfonctionnements lexicaux, 260 ont été considérés comme des erreurs, soit environ les deux tiers (66,5 %). Les autres cas (131) ont été codés comme étant des maladroresses lexicales et constituent 33,5 % du total des problèmes relevés. Le détail des types de problèmes rencontrés sera discuté plus en profondeur à la section IV-1.4.

Le contre-codage a constitué une étape importante dans la validation de notre typologie structurale des problèmes lexicaux. Si nous avons rencontré quelques problèmes à ce stade de la recherche, nous croyons que le contre-codage a permis de démontrer que notre grille d'analyse peut être utilisée par différents codeurs et conduire pourtant à des diagnostics similaires.

Cent-onze (111) problèmes, soit 28 % des 391 cas relevés lors de la correction, ont été soumis à une collègue, qui en a effectué le contre-codage. Comme nous l'avons mentionné dans notre méthodologie, aucune erreur d'orthographe lexicale n'a été incluse dans ce corpus de contre-codage, et les cas retenus l'ont été de façon aléatoire,

tant en ce qui concerne les types de problèmes que leur statut (erreur ou maladresse). De courtes explications par rapport à la typologie et à la façon de coder ont été fournies à notre collègue, mais aucune « formation pratique » à proprement parler n'a été dispensée, ce qui peut expliquer quelques surprises lors de notre première analyse du contre-codage.

1.1.1 LE CONTRE-CODAGE : PREMIÈRE ANALYSE

Lorsque nous avons effectué la comparaison des codages faits pour les 111 problèmes lexicaux, nous sommes arrivé à un accord interjuges assez peu élevé. Pour les types de problèmes lexicaux, soit les classes figurant dans notre typologie, l'accord était de 65,5 %, alors que pour les statuts des problèmes, il s'élevait à seulement 61,8 %. Les cas pour lesquels l'accord était total – même type de problème et même statut – constituaient 49,1 % des 111 cas contre-codés.

Ces divergences importantes dans le codage des problèmes lexicaux ont pu être expliquées en grande partie à la suite d'une discussion avec notre collègue qui a assuré le contre-codage. Cette conversation nous a permis de comprendre quelles ont été les sources de confusion et d'apporter quelques modifications à notre typologie.

Pour ce qui est des types d'erreurs, les divergences de codage ont pu être expliquées de diverses façons. Certains désaccords découlaient de simples erreurs d'inattention lors du codage, et le seul fait de confronter les deux codes a suffi pour qu'un des codeurs se rallie à l'avis de l'autre. D'autres cas de désaccords, moins triviaux parce que plus systématiques, nous ont par contre permis de constater certains défauts de notre typologie.

Premièrement, certaines classes d'erreurs semblaient être comprises de manière différente par les deux codeurs. Cette confusion était en fait attribuable aux noms d'origine²⁹ de certaines classes (et aux codes utilisés) ; des classes que nous aurions voulu voir s'opposer semblaient se chevaucher aux yeux de notre collègue. Par exemple, pour les problèmes relatifs au sens, nous avions à l'origine une classe « non-respect du typage des actants » et une classe « non-respect d'une composante de sens ». Comme notre contre-codeuse percevait, à raison, un problème découlant d'une

²⁹ Ces noms de classes, qui apparaissaient à l'origine dans notre typologie des problèmes lexicaux, ne figurent plus dans le présent mémoire. En effet, les confusions observées lors de la première analyse du contre-codage nous ont poussé à rebaptiser certaines classes présentées dans les sections II et III dans le but d'obtenir un travail plus clair et cohérent.

transgression du typage des actants comme un type de non-respect d'une composante de sens, nous nous sommes retrouvé avec davantage de problèmes codés « transgression d'une composante de sens » et très peu dans la classe « non-respect du typage des actants ». Le simple fait de considérer ces deux types de problèmes comme des transgressions d'une composante de sens et de rebaptiser « non-respect d'une différence spécifique » notre classe « non-respect d'une composante de sens » nous a permis d'arriver à un accord pour plusieurs cas où les codages divergeaient. Une confusion similaire a pu être observée avec les anciennes classes « régime » et « prépositions régies », puisque, encore une fois, les prépositions régies par une unité lexicale font partie intégrante de son régime. Nous avons donc remplacé ces deux classes par les suivantes : « non-respect de la structure actancielle de l'unité lexicale » et « utilisation d'une mauvaise préposition ». De nouveau, ces changements, en plus de favoriser la clarté et la cohérence de notre typologie, ont permis d'augmenter notre accord interjuges de façon importante.

Le fait de discuter des cas pour lesquels les codages différaient a aussi confirmé notre intuition selon laquelle certains problèmes peuvent être codés de plusieurs façons. En effet, si la confrontation de nos points de vue a parfois permis de convaincre l'autre codeur de se rallier à notre position, elle a surtout mis en lumière des interprétations différentes et pourtant valables pour certains problèmes. Cette constatation a eu pour effet d'augmenter les cas de double-codage, mais aussi de diminuer les désaccords dans le codage.

En ce qui concerne le statut des problèmes lexicaux, nous avons vu que l'accord interjuges n'était que de 61,8 %. Le fait de discuter avec notre collègue nous a aussi permis d'expliquer cette grande différence de perception. Nous savions d'avance que le fait d'avoir à appliquer à un problème lexical l'étiquette d'« erreur » ou de « maladresse » comportait une part de subjectivité et pouvait, par conséquent, occasionner des désaccords. Par contre, un taux de désaccord de près de 40 % nous a beaucoup surpris, et mettait en doute la pertinence de ce deuxième axe d'analyse des problèmes lexicaux.

Nous avons, dans notre méthodologie, établi certains critères pour tenter de diminuer la part de subjectivité lors de l'attribution d'un statut aux problèmes lexicaux (voir III-2.5.3), notamment le recours aux ouvrages de référence. Pour l'ensemble des propriétés de l'unité lexicale soumises à la norme prescriptive, nous avons déterminé que nous jugerions comme acceptable tout usage apparaissant dans l'un ou l'autre de

nos ouvrages de référence. Notre contre-codeuse ayant une solide expérience de lexicologue, elle s'est fiée à ses propres intuitions de linguiste plutôt que de se référer systématiquement aux ouvrages de référence. Comme les lexicologues – ou plus généralement les linguistes – adoptent souvent un point de vue plutôt descriptif et hésitent à employer des termes comme « erreur » et à condamner des usages répandus, beaucoup de problèmes ont reçu des statuts différents, alors que les ouvrages de référence auraient dû nous aider à trancher. Après une consultation commune de certains articles de dictionnaire, nous avons pu nous mettre d'accord sur la plupart des cas relevant de la norme prescriptive. Par ailleurs, le fait de réviser les statuts que nous avions attribués aux problèmes lors du codage initial a permis de noter certaines irrégularités et d'y remédier.

1.1.2 LE CONTRE-CODAGE : DEUXIÈME ANALYSE

Après avoir discuté, clarifié notre méthode de codage, révisé chacun des cas de désaccords et apporté certaines modifications à notre typologie, nous avons été en mesure de faire une deuxième analyse du contre-codage. Cette fois, le pourcentage d'accord a été beaucoup plus élevé.

Une fois les noms de nos classes d'erreurs désambiguïsés, l'accord interjuges pour les types de problèmes est passé à 86,2 %. Un pareil taux d'accord témoigne du caractère opérationnel de notre typologie. Pour ce qui est du statut des problèmes lexicaux, le taux d'accord a atteint 84,4 %, ce qui est très satisfaisant si l'on tient compte de la subjectivité inhérente au codage des problèmes relevant de dimensions qui ne sont pas soumises à la norme prescriptive. Finalement, les cas d'accord total entre les juges ont bondi de plus de 30 %, passant de 49,1 % à 80,1 %.

Comme la majorité des désaccords initiaux découlaient de confusions ayant trait aux noms des types de problèmes, nous n'avons pas cru nécessaire de faire contre-coder un autre sous-corpus une fois les noms désambiguïsés. Nous jugeons donc que nous avons atteint un taux d'accord interjuges tout à fait satisfaisant. De plus, l'étape du contre-codage a été extrêmement bénéfique dans la validation et l'amélioration de notre typologie des problèmes lexicaux.

1.2 LES ERREURS LEXICALES : RÉSULTATS

Comme nous l'avons mentionné plus haut, les 260 erreurs lexicales constituent environ les deux tiers des 391 problèmes lexicaux relevés dans notre corpus de textes, soit 66,5 %. Voici le tableau des différents types d'erreurs lexicales rencontrées, présentés du plus fréquent au moins fréquent.

Tableau VI. *Types d'erreurs lexicales rencontrées*

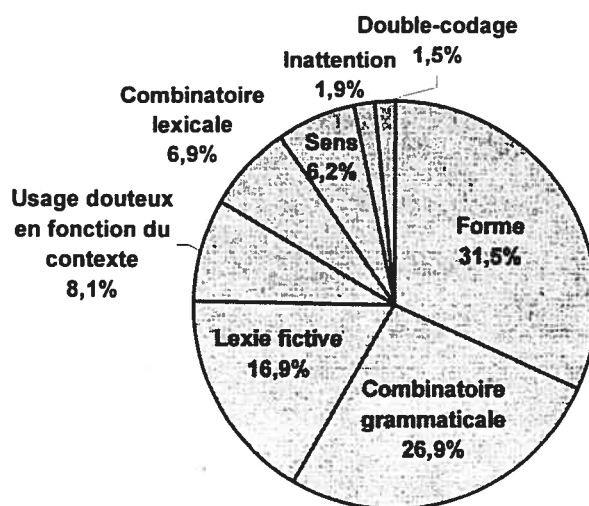
Type d'erreur lexicale	Nombre de cas	% du total des erreurs
Erreur d'orthographe lexicale	79	30,4 %
Impropriété	40	15,4 %
Non-respect de la structure actancielle de la lexie	21	8,1 %
Utilisation d'un mauvais collocatif	18	6,9 %
Utilisation d'une mauvaise préposition	17	6,5 %
Non-respect du caractère figé d'une locution	16	6,2 %
Non-respect d'une propriété inhérente à l'unité lexicale	14	5,4 %
Mélange de codes	11	4,2 %
Non-respect du typage des actants	10	3,8 %
Changement de registre de langue	7	2,7 %
Non-respect d'une différence spécifique	6	2,3 %
Erreur d'inattention	5	1,9 %
Double codage	4	1,5 %
Création d'une nouvelle acception d'un vocable	3	1,2 %
Barbarisme	3	1,2 %
Non-respect du mode verbal imposé par l'unité lexicale	2	0,8 %
Pléonasme	2	0,8 %
Mauvaise dérivation morphologique	1	0,4 %
Répétition	1	0,4 %
Utilisation d'un mauvais verbe auxiliaire	0	0,0 %
Imprécision	0	0,0 %
Cas litigieux	0	0,0 %
TOTAL	260	100,0 %

À la lecture du tableau VI, nous remarquons notamment que les erreurs d'orthographe lexicale sont, de loin, les plus présentes dans les copies d'étudiants et constituent près du tiers des erreurs (30,4 %). Le deuxième type d'erreurs lexicales le plus fréquent est l'impropriété (15,4 %). Les cinq types d'erreurs suivants – soit ceux présentant une proportion située entre 8,1 % et 5,4 % – sont tous attribuables à des transgressions de propriétés de combinatoire; ils totalisent ensemble 33,1 % des erreurs lexicales. À elles seules, les erreurs relatives au régime, soit les classes « non-respect de la structure actancielle de l'unité lexicale » et « usage d'une mauvaise préposition » représentent 14,6 % des erreurs. Les erreurs attribuables à une altération du signifié, soit par le non-respect du typage des actants (3,8 %) ou par le non-respect d'une différence spécifique

(2,3 %), constituent 6,1 % du total des erreurs. Aucun cas d'utilisation d'un mauvais auxiliaire ou d'imprécision n'a été recensé parmi les erreurs. Finalement, l'absence de cas litigieux révèle que toutes les erreurs ont pu être codées avec la typologie des problèmes lexicaux.

En regroupant les types d'erreurs en classes plus générales, correspondant aux divisions de notre typologie, nous obtenons les distributions de la figure ci-dessous.

Figure 1. Proportions des différents types d'erreurs rencontrées



Les erreurs relatives à la forme sont les plus fréquentes avec 31,5 % ; la très grande majorité de ces erreurs sont des problèmes d'orthographe lexicale, puisque seulement 3 barbarismes ont été relevés dans le corpus. La combinatoire grammaticale constitue la deuxième classe importante d'erreurs (26,9 %), alors que la combinatoire lexicale représente 6,9 % du total des erreurs. Les cas d'utilisation d'une lexie fictive s'élèvent à 16,9 % et les impropriétés sont nettement les plus fréquentes (15,4 %). Les erreurs découlant d'un usage douteux en fonction du contexte totalisent 8,1 % des cas relevés et les problèmes de sens représentent 6,2 %. Finalement, les erreurs ayant fait l'objet d'un double codage constituent 1,5 % des cas, un peu moins que les erreurs d'inattention (1,9 %).

1.3 LES MALADRESSES LEXICALES : RÉSULTATS

Des 391 problèmes codés, 131 ont reçu le statut de maladresse lexicale, ce qui représente 33,5 %. Voici le tableau des différents types de maladresses lexicales rencontrées, présentés du plus fréquent au moins fréquent.

Tableau VII. *Types de maladresses lexicales rencontrées*

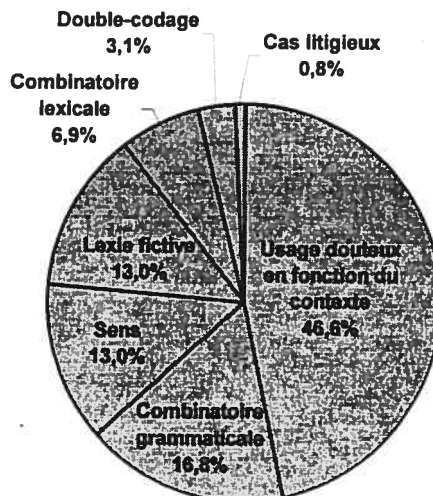
Type de maladresse lexicale	Nombre de cas	% du total des maladresses
Changement de registre de langue	28	21,4 %
Répétition	19	14,5 %
Impropriété	14	10,7 %
Non-respect d'une différence spécifique	12	9,2 %
Utilisation d'un mauvais collocatif	9	6,9 %
Non-respect du caractère figé d'une locution	8	6,1 %
Non-respect de la structure actancielle de la lexie	6	4,6 %
Mélange de codes	6	4,6 %
Pléonasme	6	4,6 %
Non-respect du typage des actants	5	3,8 %
Utilisation d'une mauvaise préposition	5	3,8 %
Double codage	4	3,1 %
Création d'une nouvelle acception d'un vocable	3	2,3 %
Non-respect d'une propriété inhérente à l'unité lexicale	3	2,3 %
Imprécision	2	1,5 %
Cas litigieux	1	0,8 %
Mauvaise dérivation morphologique	0	0,0 %
Erreur d'orthographe lexicale	0	0,0 %
Barbarisme	0	0,0 %
Non-respect du mode verbal imposé par l'unité lexicale	0	0,0 %
Utilisation d'un mauvais verbe auxiliaire	0	0,0 %
Erreur d'inattention	0	0,0 %
TOTAL	131	100,0 %

Avec 21,4 % du total des maladresses, les problèmes de registre de langue sont les plus fréquents, alors que les problèmes de répétition arrivent en deuxième position (14,5 %). L'impropriété constitue aussi un type de problème assez répandu (10,7 %), suivi de près par les maladresses découlant du non-respect d'une différence spécifique (9,2 %). Suivent les maladresses attribuables à l'utilisation d'un mauvais collocatif (6,9 %) et au non-respect du caractère figé d'une locution (6,1 %). Les autres types de problèmes représentent moins de 5 % des maladresses. Tout comme pour les erreurs, 4 cas de double codage ont été recensés parmi les maladresses. Un cas litigieux – c'est-à-dire que nous n'avons su faire correspondre à aucune de nos classes de problèmes

lexicaux – a été répertorié. Finalement, nous remarquons que six types de problèmes n'ont été associés à aucune maladresse.

La figure qui suit présente la répartition obtenue si nous regroupons les maladresses relevées selon les divisions de notre typologie.

Figure 2. *Proportions des différents types de maladresses rencontrées*



Près de la moitié des maladresses sont attribuables à un usage douteux en fonction du contexte (46,6%). Les problèmes de combinatoire grammaticale constituent le deuxième type de maladresses le plus fréquent (16,8 %), alors que les problèmes d'altération du signifié et les cas d'utilisation d'une lexie fictive représentent chacun 13 % du total des maladresses. Finalement, les problèmes liés à la combinatoire lexicale totalisent 6,9 % des maladresses.

1.4 LES PROBLÈMES LEXICAUX : RÉSULTATS GLOBAUX

Voici maintenant les différents types de problèmes rencontrés, sans distinction par rapport au statut du problème.

Tableau VIII. *Types de problèmes lexicaux rencontrés*

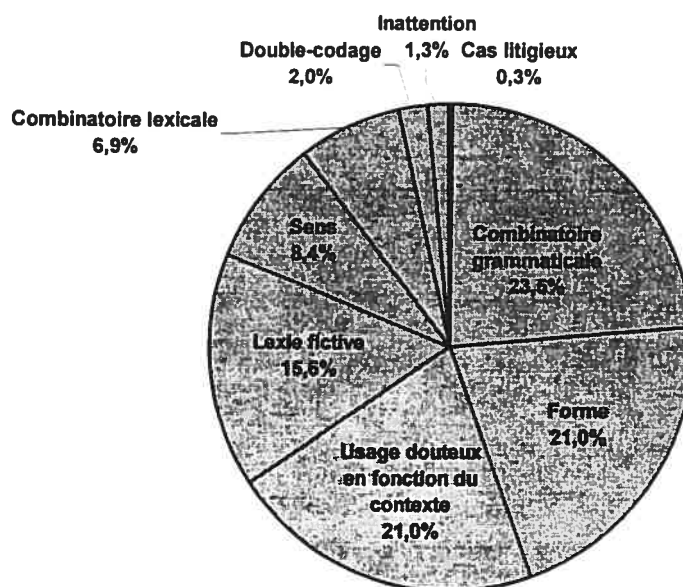
Type de problème	Nombre de cas	Nombre d'erreurs	Nombre de maladresses	% du total des problèmes
Erreur d'orthographe lexicale	79	79	0	20,2 %
Impropriété	54	40	14	13,8 %
Changement de registre de langue	35	7	28	9,0 %
Non-respect de la structure actancielle de la lexie	27	21	6	6,9 %
Utilisation d'un mauvais collocatif	27	18	9	6,9 %
Non-respect du caractère figé d'une locution	24	16	8	6,1 %
Utilisation d'une mauvaise préposition	22	17	5	5,6 %
Répétition	20	1	19	5,1 %
Non-respect d'une différence spécifique	18	6	12	4,6 %
Non-respect d'une propriété inhérente à l'unité lexicale	17	14	3	4,3 %
Mélange de codes	17	11	6	4,3 %
Non-respect du typage des actants	15	10	5	3,8 %
Pléonasme	8	2	6	2,0 %
Double codage	8	4	4	2,0 %
Création d'une nouvelle acception d'un vocable	6	3	3	1,5 %
Erreur d'inattention	5	5	0	1,3 %
Barbarisme	3	3	0	0,8 %
Non-respect du mode verbal imposé par l'unité lexicale	2	2	0	0,5 %
Imprécision	2	0	2	0,5 %
Mauvaise dérivation morphologique	1	1	0	0,3 %
Cas litigieux	1	0	1	0,3 %
Utilisation d'un mauvais verbe auxiliaire	0	0	0	0,0 %
TOTAL	391	260	131	100,0 %

L'erreur d'orthographe d'usage constitue le type le plus fréquent de problème lexical (20,2 %), suivi par l'impropriété (13,8 %). Les problèmes de registre de langue arrivent en troisième position avec 9 % du total. Les quatre types de problèmes qui suivent relèvent des propriétés de combinatoire, entre autres les deux classes concernant le régime, qui à elles seules représentent 12,5 % du total des problèmes. Parmi les classes les moins fréquentes, nous retrouvons les cas de création d'une nouvelle acception d'un vocable (1,5 %), les problèmes d'inattention (1,3 %), les barbarismes (0,8 %), les cas de non-respect du mode verbal (0,5 %), les cas d'imprécision (0,5 %) et les mauvaises dérivations morphologiques (0,3 %). Un seul problème n'a pas pu être codé avec notre typologie, ce qui représente 0,3 % des cas.

Enfin, aucun cas d'utilisation d'un mauvais verbe auxiliaire n'a été relevé dans le corpus.

En effectuant le regroupement de certains types de problèmes selon les divisions de notre grille, nous obtenons les répartitions apparaissant dans la figure ci-dessous.

Figure 3. Proportions des différents types de problèmes lexicaux rencontrés



Les problèmes relatifs aux propriétés de combinatoire grammaticale sont les plus fréquents dans notre corpus (23,5 %). Suivent les problèmes d'altération du signifiant et ceux découlant d'un usage douteux par rapport au contexte, avec 21 % dans les deux cas. Les cas d'utilisation d'une lexie fictive totalisent 15,6 % du total des dysfonctionnements observés, alors que les problèmes d'altération du signifié constituent 8,4 % du total. Finalement, les problèmes de combinatoire lexicale représentent 6,9 % des cas analysés.

2. INTERPRÉTATION DES RÉSULTATS (OBJECTIF 1)

Les pages précédentes ont révélé plusieurs informations relatives aux problèmes lexicaux des étudiants universitaires. La présente section vise quant à elle l'interprétation des résultats présentés. Nous discuterons de la fréquence de certains types de problèmes lexicaux et tenterons d'établir des tendances entre les types de problèmes et leur statut. Plutôt que de présenter séparément les erreurs et les maladresses comme nous l'avons fait précédemment, nous croyons plus intéressant de les considérer conjointement dans l'analyse de nos résultats. Cette section se terminera par une discussion sur notre typologie des problèmes lexicaux.

2.1 LES PROBLÈMES LEXICAUX : PORTRAIT GLOBAL

La première chose qui frappe à la lecture des résultats présentés au point IV-1.4 (voir figure 3) est la prépondérance des problèmes liés aux propriétés de combinatoire; en effet, si nous regroupons tous les problèmes attribuables aux propriétés de combinatoire – grammaticale (23,5 %) et lexicale (6,9 %) –, nous obtenons un total de 30,4 %. On peut donc affirmer que près du tiers des problèmes lexicaux des étudiants universitaires sont liés aux propriétés de combinatoire. Si cette observation peut sembler surprenante à première vue³⁰, certaines hypothèses nous permettent quand même de l'expliquer. Tout d'abord, nous avons vu que cette dimension de l'unité lexicale est certainement la moins connue et celle à laquelle on accorde le moins d'importance dans l'enseignement. En effet, à notre connaissance, outre le genre nominal et les parties du discours, peu de propriétés de combinatoire font l'objet d'un enseignement spécifique. De plus, les informations relatives aux propriétés de combinatoire ne sont pas toutes présentées de façon explicite et systématique dans les ouvrages de référence (par exemple, les collocations et les locutions). Enfin, nous avons aussi vu que certaines erreurs relevant de la combinatoire sont souvent considérées comme des erreurs de syntaxe ou de grammaire par les enseignants; il est donc normal que notre perspective lexicale sur ces problèmes entraîne une plus forte

³⁰ En fait, d'un point de vue linguistique, cette constatation était relativement prévisible. En tenant compte de l'importance de la combinatoire dans la maîtrise d'une unité lexicale et du nombre de propriétés qui sont impliquées, il est facile de comprendre que cette dimension puisse être une grande source d'erreurs. Par contre, ces résultats peuvent paraître étonnants dans une perspective didactique où cette dimension des unités lexicales est sans contredit la moins considérée.

proportion de problèmes lexicaux présents dans les textes. En effet, si nous avions corrigé les problèmes relevant des autres dimensions du code écrit (syntaxe, orthographe grammaticale, cohérence textuelle, etc.), notre conception « large » de l'erreur lexicale aurait forcément entraîné une baisse du nombre d'erreurs de syntaxe et d'orthographe grammaticale, puisque ces erreurs auraient été considérées comme des erreurs lexicales de combinatoire grammaticale. Nous explorerons au point IV-2.2 les différents types d'erreurs relatives aux propriétés de combinatoire.

La deuxième constatation intéressante est que les problèmes liés au signifiant (principalement des erreurs d'orthographe lexicale) ne sont pas plus fréquents que les problèmes découlant d'un usage douteux en fonction du contexte – 21 % du total dans les deux cas. Par contre, nous verrons plus loin que les problèmes liés à la forme sont à tout coup perçus comme des erreurs, alors que les problèmes liés à un usage douteux par rapport au contexte sont plus souvent considérés comme des maladresses et ne seraient pas forcément pénalisés par les enseignants.

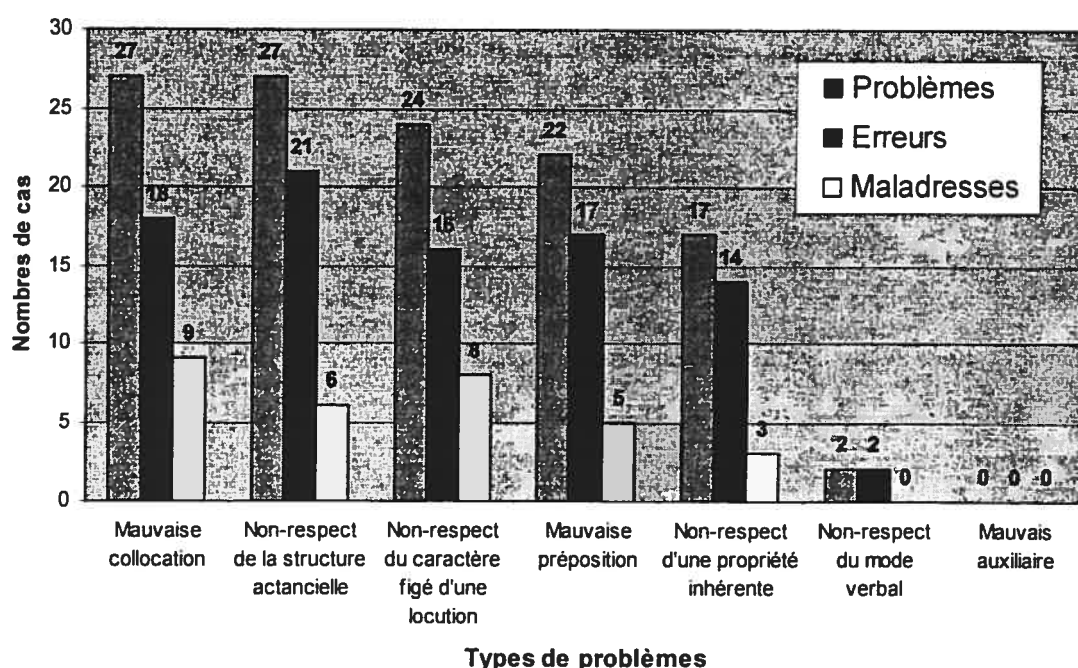
Le fait que les cas d'utilisation d'une lexie fictive constituent 15,6 % des problèmes relevés retient aussi l'attention. Nous tenterons plus bas (voir IV-2.5.1) de comprendre pourquoi ces erreurs, notamment les impropriétés, sont si fréquentes dans notre corpus. Le sens est une notion très abstraite et difficile à définir, et les problèmes relevant de cette dimension de l'unité lexicale ont sans contredit été les plus compliqués à coder. Nous expliquerons, à la section IV-2.5.1, pourquoi il est si ardu de diagnostiquer précisément le problème dans les cas relatifs à une altération du signifié.

Le point IV-2.6 sera consacré à l'étude de certains problèmes lexicaux qui ont nécessité un double codage. Nous tenterons de voir si certains types de dysfonctionnements peuvent toujours correspondre à plusieurs classes d'erreurs.

2.2 LES PROBLÈMES LEXICAUX RELATIFS À LA COMBINATOIRE

Voici tout d'abord la distribution des erreurs relatives aux propriétés de combinatoire. La colonne de gauche de chaque groupe représente le nombre de cas du type concerné, alors que celle du centre et celle de droite indiquent respectivement le nombre d'erreurs et le nombre de maladresses de ce type.

Figure 4. Les problèmes de combinatoire



2.2.1 LA COMBINATOIRE LEXICALE

Les problèmes relatifs aux propriétés de combinatoire lexicale constituent 6,9 % (27 cas) des problèmes relevés dans le corpus. La grande majorité des problèmes sont des collocations ayant un nom comme base et un verbe comme collocatif, par exemple :

- (1) L'élève devrait être plus pénalisé ou bien voir **arriver des conséquences** directes sur son apprentissage... [sujet A-01]
- (2) Est-ce parce que les enseignants ne savent plus **employer leur autorité**? [sujet A-19]
- (3) ... les parents devraient **avoir** beaucoup **plus de discipline** à la maison avec les jeunes. [sujet K-04]
- (4) Elle doit également travailler à son évolution tout en **appliquant un plus grand respect** de soi et d'autres personnes. [sujet K-01]

Il y a aussi quelques cas de collocations ayant un verbe comme base et un adverbe comme collocatif (nos bouches se sont **grandement ouvertes** et nos yeux se sont **vigoureusement agrandis**. [sujet H-22]). On remarque que les problèmes de collocations, bien qu'ils soient soumis à la norme descriptive, ont été perçus comme des erreurs deux fois plus souvent que des maladresses. Ceci semble indiquer que dans plusieurs cas, il est possible d'avoir une intuition assez forte sur l'acceptabilité d'une collocation.

2.2.2 LE RÉGIME

À elles seules, les erreurs relatives au régime, soit les classes « non-respect de la structure actancielle de l'unité lexicale » et « usage d'une mauvaise préposition », représentent 14,6 % des erreurs et 12,5 % de tous les problèmes observés. Pourtant, plusieurs informations relatives au régime apparaissent dans les ouvrages de référence et de telles erreurs pourraient facilement être évitées.

Les problèmes de non-respect de la structure actancielle de l'unité lexicale représentent 6,9 % du total (27 cas). La totalité des problèmes de ce type, sauf deux, mettent en jeu des verbes, comme les exemples suivants :

- (5) Un jour, c'est **à mon frère** Julien que *j'avais* décidé (sic) de **donner l'accès**. [sujet H-20]
- (6) **Elle lui informe ses déplacements** pour le besoin de l'enquête. [sujet G-18]
- (7) Cette péripétie se terminera très mal et **Carter récoltera d'une nuit** en prison. [sujet G-02]

Les cas où il manque un des compléments, comme l'exemple (5) dans lequel il manque le complément du nom ACCÈS, sont assez fréquents. Il y a aussi plusieurs cas pour lesquels le dysfonctionnement semble découler d'une contamination d'un verbe quasi-synonyme qui présente une certaine parenté phonologique. Par exemple, le verbe INFORMER de (6) possède la structure du verbe INDiquer, et la microsyntaxe du verbe RÉCOLTER de (7) est construite sur le modèle d'ÉCOPER.

Pour ce qui est des cas d'utilisation d'une mauvaise préposition, ils constituent 5,6 % du total (22 cas) et présentent une répartition assez égale entre les prépositions régies par des verbes, des adjectifs et des noms. Voici des exemples de cas :

- (8) Cependant, une discipline trop rigide pourrait affecter l'**intérêt** des jeunes **envers** une certaine activité... [sujet K-10]
- (9) Par exemple, il n'est pas rare d'entendre qu'un élève a été **impoli face à** son professeur. [sujet A-14]
- (10) Le réalisateur Alfonso Cuarón **a réussi**, avec une main de maître (sic), **de** recréer le monde fantastique du livre... [sujet G-12]

Nous remarquons finalement que la majeure partie des problèmes relatifs au régime sont perçus comme des erreurs (38 cas sur 49), probablement à cause du fait que le régime est soumis la norme prescriptive et qu'il est généralement possible de trouver ces informations dans le dictionnaire.

2.2.3 LE CARACTÈRE FIGÉ DES LOCUTIONS

Le non-respect du caractère figé des locutions constitue le troisième type d'erreurs relatives aux propriétés de combinatoire. Cette classe de problèmes lexicaux est un peu floue, puisqu'il existe de nombreux types de transgressions du caractère figé d'une locution. Voici deux exemples pour lesquels le dysfonctionnement semble être attribuable au fait d'avoir fléchi la locution selon les règles de grammaire, alors qu'elle aurait dû demeurer invariable :

- (11) Les fiançailles se firent très rapidement [...] et les noces se fêtèrent **en grande**. [sujet G-01]
- (12) ... les élèves sont assez matures [...] pour suivre le cours en étant à **leurs affaires**. [sujet A-04]

Certains cas présentaient plutôt une déformation de la locution par le remplacement d'un des termes constituants par une autre unité lexicale, ou le simple ajout d'un élément.

- (13) ...il n'est pas question d'une discipline à **deux temps**. [sujet A-06]
- (14) ...une scène d'action à **en couper le souffle**. [sujet G-32]
- (15) Le réalisateur Alfonso Cuarón a réussi, **avec une main de maître**, de (sic) recréer le monde fantastique du livre... [sujet G-12]

Les cas les plus spectaculaires de déformation des locutions sont sans doute les télescopages d'expressions, comme dans les exemples suivants :

- (16) Les vedettes, Tom Hanks et Meg Ryan, se retrouvant encore une fois sur **le plateau du grand écran**... [sujet G-17]
- (17) Il ne reste qu'à espérer que ceux qui enseignent dans **le juste et bon milieu** deviennent la majorité. [sujet A-09]
- (18) ... cela leur permet de **prendre la responsabilité de leurs études entre leurs mains**. [sujet A-04]

2.2.4 LES PROPRIÉTÉS INHÉRENTES À L'UNITÉ LEXICALE

Cette classe d'erreurs lexicales aurait plutôt dû porter le nom de « non-respect d'autres propriétés inhérentes à l'unité lexicale », puisque le régime aussi une propriété inhérente dont la typologie rend compte par d'autres classes. Mais au-delà de cette remarque taxinomique, quelles étaient les erreurs associées à cette classe?

Nous avons relevé 17 problèmes lexicaux de ce type, dont 14 ont été perçues comme des erreurs. Les problèmes les plus fréquents étaient ceux ayant trait à l'invariabilité d'une unité lexicale, comme les exemples suivants :

- (19) Amateurs d'**actions**, vous ne serez pas déçus car ce film en regorge. [sujet G-32]
- (20) Yassin, quant à lui, il ne représente (sic) qu'un petit enfant dont **les qualités de vie** ont été brimées par sa surdité. [sujet G-08]

Certaines erreurs découlaient plutôt du non-respect de propriétés comme le genre nominal ou la partie du discours :

- (21) Puisque l'enfant à (sic) tendance à s'associer à **différents idoles** comme point de repère... [sujet K-11]
- (22) Étrangement, les **victimes** ont **tous** un point en commun. **Ils révèlent** êtres (sic) ses anciens amants. [sujet G-18]
- (23) Jessica est **alcoolisme**. [sujet G-18]
- (24) Cet endroit était idéal pour lire **tranquille**... [sujet H-12]

Finalement, les quelques problèmes qui ne correspondent pas à ceux mentionnés plus haut (invariabilité, genre nominal ou partie du discours) sont ceux qui ont été perçus comme des maladroites. La seule erreur qui a posé problème et pour laquelle il y aurait peut-être lieu de créer une classe est l'usage non pronominal de RÉVÉLER dans (22), alors qu'il aurait dû être utilisé pronominalement. Par contre, un seul cas semblable a été relevé dans le corpus.

2.2.5 LE MODE VERBAL IMPOSÉ PAR L'UNITÉ LEXICALE

Deux cas de ce type ont été repérés :

- (25) **Bien que** le fouet **n'est pas utilisés** (sic), certaines écoles... [sujet A-12]
- (26) Elle se débrouille donc pour l'éloigner, en **espérant que** cette jeune femme **ait** une meilleure vie de (sic) ce qu'elle aurait vécu au couvent. [sujet G-21]

Dans les deux cas, les problèmes ont reçu le statut d'erreur, puisque des informations relatives à cette propriété apparaissent dans les ouvrages de référence.

2.2.6 L'UTILISATION D'UN MAUVAIS VERBE AUXILIAIRE

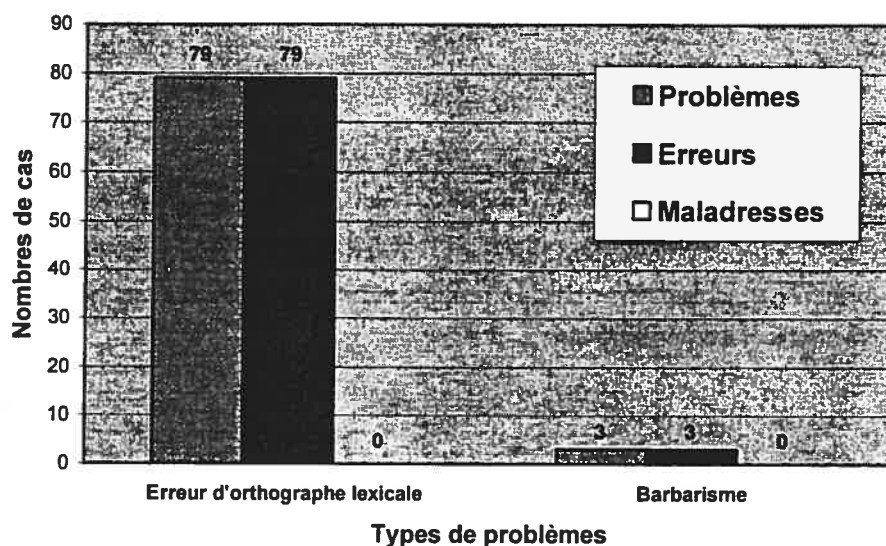
Aucun cas d'utilisation d'un mauvais verbe auxiliaire n'a été repéré dans notre corpus, ce qui nous a beaucoup étonné. Nous croyons quand même qu'il est nécessaire de

garder cette classe d'erreur dans notre typologie, puisque nous sommes convaincu que ce problème risque d'être rencontré dans des copies d'élèves ou d'étudiants. Afin d'obtenir quelques commentaires de la part des étudiants sur les problèmes de ce type, nous en avons inclus un cas dans la seconde partie du test de notre deuxième objectif.

2.3 LES PROBLÈMES LEXICAUX RELATIFS À LA FORME

Voici un schéma présentant le nombre de problèmes relatifs à la forme et leur répartition selon leur statut :

Figure 5. *Les problèmes de forme*



Ce qui frappe d'abord à la vue de la figure 5 est l'absence complète de maladresses parmi les 82 cas de problèmes découlant d'une altération du signifiant. Par contre, ce fait n'est pas vraiment étonnant si l'on considère que la forme d'une unité lexicale est sa dimension la plus soumise à la norme prescriptive. En effet, la rigidité en ce qui concerne l'orthographe lexicale explique que l'ensemble des problèmes de cette nature aient reçu le statut d'erreur.

Il est tout de même surprenant de constater la présence d'autant de problèmes relatifs à la forme, alors que les textes du corpus ont été rédigés à la maison, que les étudiants avaient accès aux ouvrages de référence (et même au vérificateur orthographique dans plusieurs cas, puisque la plupart des travaux étaient faits à l'ordinateur) et que l'orthographe d'un mot-forme est sa dimension la plus simple à vérifier dans le

dictionnaire. Il y a donc lieu de supposer que les étudiants n'ont pas eu recours aux ouvrages de référence pour vérifier l'orthographe. Est-ce parce qu'ils ne doutent pas de leur maîtrise de l'orthographe lexicale, parce qu'ils savent qu'un problème de forme ne compromettra pas la compréhension de leur texte ou tout simplement parce qu'ils sont trop préoccupés par les autres dimensions du code écrit? À leur décharge, nous pouvons tout de même affirmer que l'orthographe des locutions peut être plus difficile à vérifier, puisque ces dernières sont souvent enchâssées à l'intérieur de l'article d'une des unités lexicales la composant.

2.3.1 LES ERREURS D'ORTHOGRAPHE

Nous remarquons dans la figure 5 la prépondérance des erreurs d'orthographe lexicale, qui constituent à elles seules 20,2 % du total des problèmes lexicaux relevés. Nous nous attendions bien sûr à une forte présence d'erreurs d'orthographe lexicale, puisque ce type d'erreurs avait déjà été recensé dans les recherches mentionnées dans notre problématique (Roy *et al.*, 1995 : 30 ; Lefrançois, 2004). Parmi les erreurs d'orthographe³¹, nous avons remarqué de nombreux cas relatifs à l'emploi du trait d'union (*bien être* [sujet K-01], *sous sol* [sujet H-23], *non-seulement* [sujet G-29], *papier-peint* [sujet H-02], etc.). Certains problèmes ont trait à la division des mots-formes (*dégustation de guimauves au tour du feu* [sujet H-25], *vivre au près de ma grand-mère* [sujet H-18], *il y avait une grosse couverture parterre* [sujet H-02], *pour en savoir d'avantage* [sujet G-20], etc.) et d'autres à l'emploi des majuscules (*partir dans le nord* [sujet G-19], *l'Est de la Pologne* [sujet G-11], *la marine Allemande* [sujet G-21], etc.). Les locutions semblent aussi constituer une source importante d'erreurs d'orthographe (*entre autre* [sujet H-10], *de qualités* [sujet H-08], *ils font partis* [sujet G-12], *au dépend de* [sujet G-02], *peut importe* [sujet A-06], *quoiqu'il en soit* [sujet A-04], etc.).

³¹ Nous ne faisons ressortir ici que quelques observations relatives aux erreurs d'orthographe, nous ne prétendons en aucun cas en proposer une catégorisation complète. Catach (1995) propose une typologie exhaustive de ce type d'erreurs.

2.3.2 LES BARBARISMES

Voici les trois barbarismes que nous avons relevés dans notre corpus :

(27) ... le président Bush perd la vedette **au profil de** l'absurdité humaine. [sujet G-26]

(28) **Tant qu'à moi**, j'ai eu la chance d'avoir un endroit où je pouvais profiter de la nature... [sujet H-11]

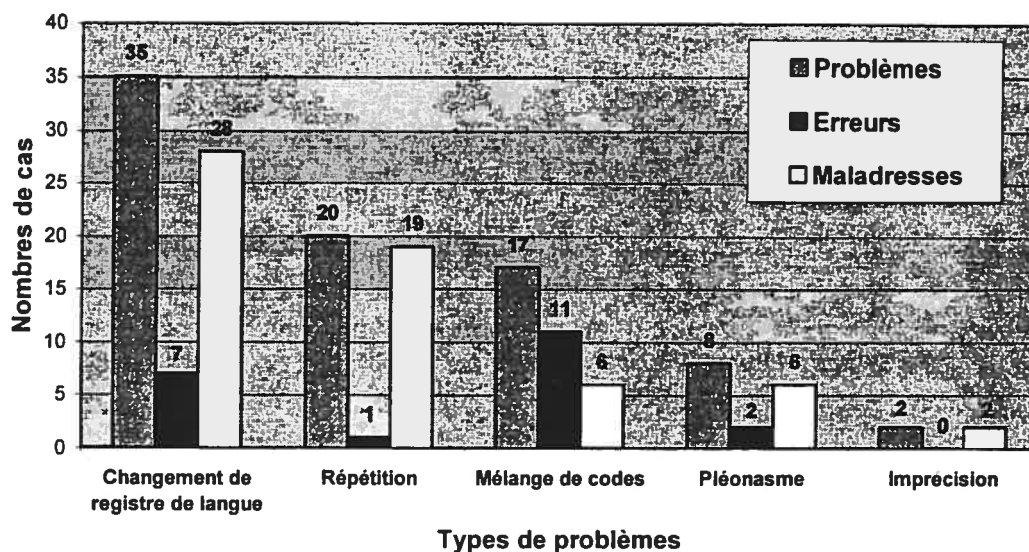
(29) ... les méthodes de punition et de **félécitation**... [sujet K-14]

Ces trois cas de barbarismes sont somme toute assez différents. L'exemple (27) présente une locution complètement déformée. Il est assez surprenant que le sujet G-26 « entende » *profil* dans cette expression et n'ait pas compris que le terme PROFIT, qui garde son sens à l'intérieur de la locution AU PROFIT DE, conviendrait beaucoup mieux. Dans le cas (28), la confusion entre **tant qu'à moi* et QUANT À MOI semble beaucoup plus facile à expliquer. En effet, ce barbarisme est très répandu et constitue presque la norme à l'oral. Par conséquent, nous pouvons comprendre que cette forme entre en concurrence avec la forme correcte chez certaines personnes, d'autant plus que l'expression TANT QU'À, bien que familière, existe vraiment (*Tant qu'à attendre demain, aussi bien partir maintenant.*). Finalement, le cas (29) est un cas hybride entre le barbarisme et l'erreur d'orthographe d'usage, puisque la différence est à peine perceptible à l'oral, l'erreur se trouvant à l'intérieur d'une syllabe non accentuée. Nous pouvons donc supposer qu'il existe plusieurs types de barbarismes, mais notre corpus limité ne nous permet pas d'en établir ici une liste exhaustive.

2.4 LES PROBLÈMES LEXICAUX DÉCOULANT D'UN USAGE DOUTEUX PAR RAPPORT AU CONTEXTE

Voici la distribution des problèmes découlant d'un usage douteux par rapport au contexte :

Figure 6. Les problèmes découlant d'un usage douteux en fonction du contexte



L'analyse de la figure 6 nous permet d'abord de constater que la très grande majorité des problèmes relatifs au contexte ont été perçus comme des maladresses (61 cas sur 82), sauf dans le cas des mélanges de codes. Nous avons mentionné que les problèmes relatifs à un usage douteux d'une unité lexicale en fonction du contexte laissaient place à plus de subjectivité, puisqu'ils sont en quelque sorte soumis à une norme « stylistique » plutôt floue. C'est probablement pour cette raison que ces problèmes ont plutôt été considérés comme des maladresses. La seule classe pour laquelle le nombre d'erreurs est plus élevé est « mélange de codes ». Cette classe contient majoritairement des anglicismes, sur lesquels les ouvrages de référence sont généralement assez clairs.

2.4.1 LES PROBLÈMES DE REGISTRES DE LANGUE

Les problèmes d'utilisation d'un registre de langue inapproprié sont les plus fréquents avec 9,0 % du total de tous les problèmes. La très grande majorité de ces problèmes ont été perçus comme des maladresses (28 sur 35). Nous avons tenté de comprendre

pourquoi certains problèmes ont plutôt reçu le statut d'erreur. Il semble que les unités lexicales appartenant à un niveau de langue négligé ou carrément vulgaire sont perçues comme erronées dans le cadre d'un texte scolaire, alors que celle relevant du langage parlé et du langage familier sont plus difficilement condamnables. Par exemple, les unités lexicales CRAMER [sujet H-18], SE FOUTRE (de) [sujet A-18] et À CAUSE QUE [sujet G-22] ont reçu le statut d'erreur, alors que les lexies TANNÉ_{Adj} [sujet G-31], QUESTION DE [sujet G-34], CHIGNER [sujet A-18] et VISITE [*de la vaisselle réservée à la grande visite*] [sujet H-20] font partie des maladresses. Parmi les maladresses, nous avons aussi remarqué l'utilisation intransitive de certains verbes transitifs comme ANGOISSER [sujet G-08] et DÉPRIMER [sujet G-19]. Quelques rares cas d'utilisation d'unités lexicales appartenant à un niveau de langue trop littéraire ou vieilli produisaient aussi une impression de cassure dans l'uniformité du registre de langue employé. Nous pensons par exemple à l'utilisation de CEPENDANT [sujet G-09] au sens de 'pendant ce temps', qui siérait davantage à un roman de Zola qu'au maladroït résumé du film *Danse lascive 2* qui le contenait.

2.4.2 LES RÉPÉTITIONS

Les cas de répétition sont aussi assez fréquents dans le corpus : 5,1 % de l'ensemble des problèmes. Par contre, la totalité des problèmes de répétition, sauf un, ont reçu le statut de maladresse. Le seul cas qui ait reçu le statut d'erreur présentait onze occurrences du mot-forme *endroit* dans un texte de 272 mots, dont cinq avec le modificateur MERVEILLEUX. La plupart des autres cas sont des répétitions « en écho » comme les exemples suivants :

- (30) Le **dernier** film qui m'a le plus marqué **dernièrement**, c'est celui du réalisateur Micheal Moore : Fahrenheit 9/11. [sujet G-26]
- (31) Il y en a certains que je connais par cœur tellement **je les ai regardés souvent**. Encore aujourd'hui, il m'arrive **souvent d'en regarder**... [sujet H-05]
- (32) Cet endroit **se situait** dans le creux de la garde-robe **située** à côté de ma chambre. [sujet H-12]

Il serait difficile pour un enseignant de vraiment pénaliser de tels cas de répétition, mais les souligner en demandant à l'étudiant de tenter de reformuler sa phrase de sorte à éliminer la répétition pourrait sensibiliser ce dernier à ce problème et le motiver à varier le vocabulaire de ses textes.

2.4.3 LES MÉLANGES DE CODES

Les 17 cas de mélange de codes relevés dans le corpus sont des anglicismes, comme nous pouvions nous y attendre. La plupart ont reçu le statut d'erreur, puisqu'ils étaient condamnés par les ouvrages de référence, comme les anglicismes suivants :

- (33) ... ce qui permet aux enseignants de **mettre** beaucoup **d'emphase** sur la qualité de l'enseignement... [sujet A-10]
- (34) Mais peut-on affirmer que les jeunes ne saisissent pas assez les **opportunités** de s'amuser en plein air? [sujet H-11]
- (35) Le film **prend place** dans la localité de Beaufort High. [sujet G-29]

Parmi les cas d'anglicismes ayant été considérés comme des maladresses, nous retrouvons entre autres le nom FUTUR employé au sens d'« avenir » (4 occurrences), que les ouvrages de référence qualifient d'abusif sans parler d'anglicisme.

2.4.4 LES PLÉONASMES

Huit cas de pléonasme ont été trouvés dans notre corpus et seulement deux d'entre eux ont été perçus comme des erreurs. Ces deux cas sont :

- (36) Plus le temps a passé, **plus** les générations sont devenues **davantage** permissives. [sujet A-02]
- (37) ... une jeune fille qui voulait changer le monde en envoyant, **comme par exemple**, de la nourriture dans... [sujet G-09]

Les autres cas d'utilisation de pléonasmes donnaient lieu à des phrases dont le sens n'était pas toujours clair. On peut donc supposer que les étudiants ont écrit ces phrases sans trop s'interroger sur leur sens. Les exemples suivants – en particulier (40) – témoignent du caractère opaque de ces phrases.

- (38) De plus, l'**importance** du lien entre l'étudiant et son enseignant **est essentiel** (sic). [sujet A-13]
- (39) ... celle-ci **s'enferme** dans sa garde-robe **afin de s'y réfugier**. [sujet G-06]
- (40) Il y a un **mystère** dans cette **énigme**. [sujet G-18]

2.4.5 LES CAS D'IMPRÉCISION

Seulement deux cas d'imprécision ont été recensés. Cette classe est sans doute la plus subjective de toutes et les deux cas relevés l'ont été pour des raisons purement stylistiques; il semble que les phrases (41) et (42) contiennent trop de « mots » pour ce

qu'elles veulent dire et paraissent très maladroites. En fait, il est assez étonnant que des francophones aient produit ces tournures de phrase.

(41) Ce livre se déroule **dans les années actuelles**... [sujet G-16]

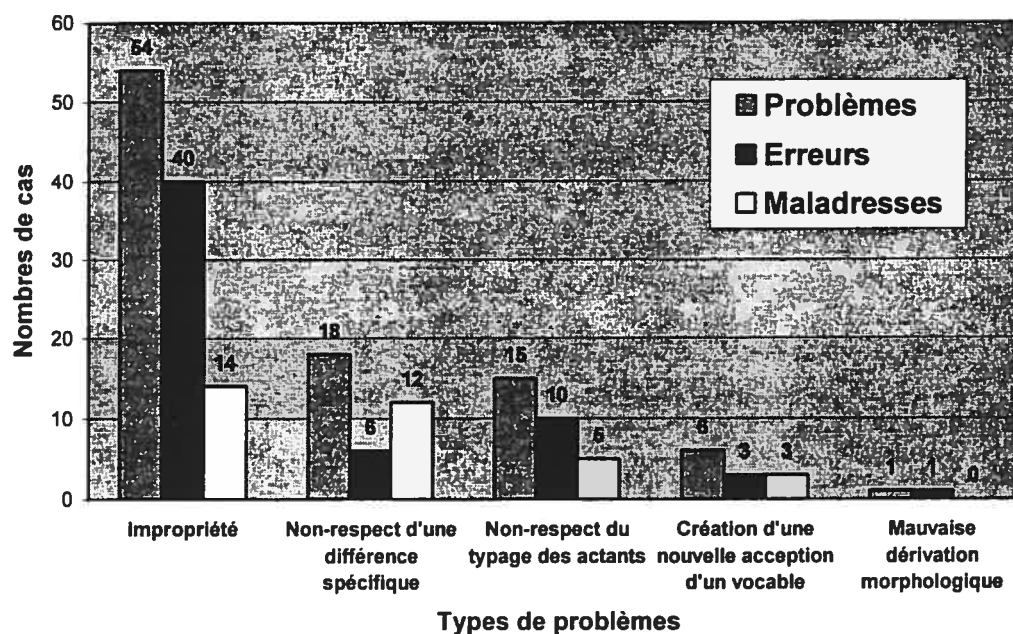
(42) ... au pays des merveilles : le parc qui longeait **ma rue de résidence**. [sujet H-21]

Il semble qu'il aurait semblé beaucoup plus naturel d'employer *de nos jours* en (41) et *ma rue* en (42). Par contre, ce jugement est purement subjectif et c'est pour cette raison que nous parlons plutôt de maladresses dans ces cas-ci.

2.5 LES CAS D'UTILISATION D'UNE LEXIE FICTIVE ET LES PROBLÈMES DE SENS

La figure qui suit présente la distribution des différents types d'utilisation d'une unité lexicale fictive selon leur statut, ainsi que les problèmes relatifs à l'altération du signifié.

Figure 7. Les cas d'utilisation d'une lexie fictive et les problèmes de sens



L'improprété est clairement le type d'utilisation d'une lexie fictive le plus fréquent, avec 54 cas relevés. Les autres problèmes de sens sont beaucoup moins répandus. Nous tenterons d'expliquer, à la section IV-2.5.1, pourquoi il faut selon nous

relativiser ces résultats concernant les problèmes de sens. Nous observerons ensuite les deux autres cas d'utilisation d'une lexie fictive : la création d'une nouvelle acception d'un vocable et les mauvaises dérivations morphologiques.

2.5.1 LES IMPROPRIÉTÉS ET LES AUTRES PROBLÈMES DE SENS

De tous les types de problèmes lexicaux relevés dans le corpus, le deuxième plus fréquent est l'impropriété (13,8 %). Cette observation nous laisse supposer que les étudiants ne remettent pas en doute leur compréhension du sens des unités lexicales ou encore qu'ils utilisent peu, ou mal, le dictionnaire. Notre deuxième objectif nous permettra de vérifier ces hypothèses.

En ce qui concerne les problèmes attribuables à l'altération du signifié, soit par le non-respect d'une différence spécifique (4,6 %) ou par le non-respect du typage des actants (3,8 %), nous remarquons qu'ils sont beaucoup moins fréquents que les impropriétés, qui constituent pourtant des dysfonctionnements plus graves, puisque le sens exprimé par l'unité lexicale employée ne correspond pas au sens visé. Il faut cependant relativiser un peu ces résultats par rapport aux problèmes de sens, puisque la classification de ce type d'erreur dans les trois classes de notre typologie s'est avérée extrêmement ardue, notamment à cause de la difficulté à cerner exactement la composante de sens qui a été transgressée. Il se peut donc que plusieurs problèmes aient été codés « impropriété » faute de pouvoir déterminer avec précision la dimension du sens transgressée. La classe « impropriété » a donc en quelque sorte servi de classe fourre-tout aux problèmes de sens qui nous laissaient perplexe.

Nous tenterons de démontrer, par des exemples, la difficulté de classer des problèmes relatifs au sens. Prenons d'abord des cas tranchés :

(43) Pour ne pas nous faire repérer par de méchants intrus, nous **arborions** un habit de camouflage brun-vert. [sujet H-04]

(44) **Pour en terminer avec** mon endroit merveilleux, je dois ajouter qu'il était rempli de plusieurs jouets... [sujet H-14]

Les phrases (43) et (44) présentent clairement des cas de non-respect d'une ou plusieurs différences spécifiques. Dans (43), le verbe ARBORER a été utilisé pour exprimer 'porter', mais sans tenir compte du fait qu'il véhicule aussi l'idée de 'porter ostensiblement, avec fierté'. Évidemment, ces composantes de sens sont incompatibles avec le contexte, puisque l'auteur mentionne que les habits de camouflage étaient

portés précisément pour ne pas attirer l'attention! Dans (44), l'utilisation de la locution EN TERMINER (avec) pour exprimer 'pour conclure' pose problème, puisque le sujet H-14 n'a pas tenu compte du fait que cette expression a une connotation particulière. L'expression EN TERMINER (avec quelque chose) signifie 'avoir enfin fini avec quelque chose', comme si la personne qui l'employait en avait vraiment assez de cette question. Évidemment, cette expression ne convient pas du tout pour annoncer la conclusion d'un texte qui porte sur un endroit merveilleux et produit un effet plutôt comique. Les phrases (45) et (46) contiennent des exemples clairs de non-respect du typage des actants :

(45) Des multitudes d'*objets* y *étaient emmenés*. [sujet H-18]

(46) ... elle recevra une *poudre magique* sur la tête qui lui *accordera* son vœux (sic). [sujet G-06]

Dans le cas de la phrase (45), il y a incompatibilité entre le verbe EMMENER et le complément OBJET³², puisque ce verbe exige que son complément dénote un être animé. Dans (46), c'est plutôt le sujet du verbe qui pose problème. En effet, une poudre ne peut pas 'consentir à donner ou à réaliser' un vœu; le verbe ACCORDER demande un sujet qui ait la capacité de consentir, donc de réfléchir.

Voici maintenant des cas clairs d'impropriétés, c'est-à-dire d'unités lexicales employées pour exprimer un sens qui ne leur correspond pas.

(47) Juste des petits mots d'encouragement dans une classe peuvent faire une très grande différence *en tant que* réussite scolaire. [sujet K-04]

(48) Puisque l'enfant à (sic) tendance à *s'associer* à différents (sic) idoles comme point de repère... [sujet K-11]

(49) Par contre, je crois que la discipline est souvent mal comprise par les élèves un peu plus *expressifs* et s'en trouve donc plus contestée... [sujet K-09]

(50) ... inspiré d'une histoire véridique : l'*invasion* d'une centaines de juifs d'un camp de la mort... [sujet G-11]

Les quatre cas présentés contiennent clairement des impropriétés, puisque les sens véhiculés par les unités employées ne correspondent en rien aux sens visés. Dans (47), la locution EN TANT QUE, qui signifie 'considéré comme', tente d'exprimer 'dans'. Par le verbe S'ASSOCIER de (48), l'auteur voulait exprimer le sens 's'identifier'. Le terme EXPRESSIF de la phrase (49) est utilisé pour signifier 'turbulent', alors qu'aucun lien sémantique n'existe entre ces deux adjectifs. Finalement, la phrase (50) présente une confusion entre les paronymes INVASION et ÉVASION. Nous discuterons à la section IV-

³² Bien entendu, le verbe EMMENER étant ici utilisé à la forme passive, le complément direct *objets* joue ici le rôle syntaxique de sujet.

2.9.1 de l'ajout de sous-classes à l'intérieur des impropriétés afin d'en distinguer les différents types.

Prenons maintenant un cas moins clair et tentons de comprendre pourquoi il est difficile de diagnostiquer certains problèmes reliés au sens.

(51) Ces 92 pages nous **récitent** les péripéties d'une seule journée dans la vie de Dastaguir... [sujet G-08]

Dans (51), il est évident qu'il y a un problème de sens avec le verbe RÉCITER. Si on regarde dans le *Petit Robert électronique*, nous trouvons trois acceptions du verbe RÉCITER : 'dire à haute voix', 'dire sans sincérité ni véracité' et 'chanter un récitatif'. Comme aucune de ces acceptions ne correspond à l'utilisation de RÉCITER présente dans (51), il faut considérer qu'il s'agit d'une impropriété, ce qui a été notre position. Pourtant, le choix du verbe RÉCITER ne semble pas accidentel dans ce cas-ci et nous aurions intuitivement été tenté de dire que c'est plutôt une composante de sens qui a été transgressée. Mais laquelle? Ni 'dire' ni 'à voix haute' ne fonctionnent dans le contexte de la phrase (51). Nous pourrions aussi affirmer qu'il s'agit d'un problème de typage des actants, puisque c'est la cooccurrence du sujet *ces pages* avec le verbe RÉCITER qui pose problème, puisque les pages ne peuvent pas dire quelque chose à voix haute. Finalement, il y aurait lieu de penser que l'étudiant voulait simplement utiliser le verbe RELATER, dont la forme est assez proche du verbe choisi... Il y a donc quatre interprétations possibles pour le problème présent dans (51), et cela sans compter l'hypothèse selon laquelle l'auteur a dérivé un verbe RÉCITER correspondant au nom RÉCIT, qui peut dénoter aussi bien un texte oral qu'écrit.

La phrase (52) contient un autre cas problématique :

(52) Usant d'ingéniosité, Jessica reconnecte les fils du téléphone le mieux possible et, après maintes **reprises**, elle réussit à joindre un jeune homme... [sujet G-24]

L'unité lexicale REPRISE a ici été utilisée pour exprimer le sens 'tentative'. Aucune des acceptions du vocable REPRISE ne correspond au sens 'tentative'. Malgré cela, l'utilisation du terme REPRISE ne semble pas si maladroite dans (52), ce qui pourrait nous laisser croire que seule une composante de sens distingue REPRISE de TENTATIVE. Pourtant, ce n'est pas le cas et nous n'avons pas réussi à le prouver en consultant les ouvrages de référence. En fait, il y a lieu de supposer que le sujet G-24 voulait exprimer 'plusieurs tentatives', mais que le choix de l'adjectif MAINT pour exprimer 'plusieurs' a fait « dérailler » le choix de TENTATIVE vers REPRISE par une contamination

de la locution À MAINTES REPRISES. Cette erreur constituerait donc une sorte d' « ellipse sémantique » ; si *Jessica a effectué plusieurs tentatives*, c'est qu'elle *a tenté à maintes reprises*.

Avec ces deux cas, nous avons démontré qu'il a parfois fallu considérer certains problèmes comme des impropriétés alors que notre intuition nous aurait plutôt poussé à parler de non-respect d'une composante de sens. Nous voyons deux explications au fait que les problèmes de sens soient si difficiles à diagnostiquer.

Premièrement, ils ne sont détectés que par rapport à une incompatibilité sémantique avec des éléments du contexte. Par conséquent, par une seule occurrence d'une lexie dans un seul contexte, nous devons tenter de définir le sens qu'a l'unité lexicale en question pour la personne qui l'a utilisée. Une fois que nous avons reconstruit le sens que la personne voulait exprimer par inférence par rapport au contexte, il faut comparer ce sens visé au sens réellement véhiculé par l'unité lexicale en jeu. Plutôt que d'aller du sens vers le texte, comme le favorise la théorie sur laquelle nous nous appuyons, nous devons reconstituer le sens à partir d'un échantillon de texte. Il est donc normal que les classes d'erreurs relatives au sens que nous avons déterminées d'un point de vue purement théorique soient moins bien définies lorsque l'on procède à contre-courant. À cet effet, Polguère (à paraître) souligne d'ailleurs que, si les théories linguistiques permettent de décrire de façon élégante la langue, elles s'avèrent moins efficace pour décrire la parole³³ avec toutes les imperfections qu'elle présente.

« La lexicologie structurale synchronique, du type explicative et combinatoire, a développé au cours des années des outils puissants pour modéliser les phénomènes lexicaux. Il faut maintenant que de tels outils, avant tout destinés à la modélisation d'un modèle statique idéalisé, soient adaptés et utilisés pour décrire la « vraie » langue : celle que l'on parle. » (Polguère, à paraître : 21)

La deuxième raison pour laquelle il est difficile de spécifier quelle composante de sens a été transgressée est le fait que les dictionnaires commerciaux que nous utilisons comme ouvrages de référence ne fournissent pas systématiquement de décompositions sémantiques complètes et structurées comme celles élaborées dans le cadre de la théorie Sens-Texte (voir II-1.4). Par conséquent, certaines informations, par exemple le typage des actants, peuvent ne pas apparaître dans tous les cas. Il est donc pratiquement impossible de cibler spécifiquement la composante de sens qui fait

³³ Revoir opposition *langue* ~ *parole* à la section II-1.2.1.

défaut, alors que les composantes de sens ne sont pas toutes présentes dans les définitions.

Nous croyons donc que nos classes d'erreurs relatives à une altération du signifié sont valides théoriquement et qu'elles peuvent servir à caractériser certains problèmes lexicaux clairs. Par contre, nous sommes tout à fait conscient des difficultés rencontrées lors de l'analyse de plusieurs problèmes ayant trait au sens. L'ajout de certaines sous-classes parmi les impropriétés devrait nous permettre de gagner en précision dans la caractérisation des problèmes de sens, même si nous croyons qu'il y aura toujours une classe résiduelle pour les problèmes de sens pour lesquels nous pouvons difficilement établir un diagnostic précis.

2.5.2 LES CAS DE CRÉATION D'UNE NOUVELLE ACCEPTION D'UN VOCABLE

Seulement six cas de création d'une nouvelle acception d'un vocable ont été relevés dans notre corpus. De ces six cas, quatre étaient de nouvelles acceptions d'un vocable créées sur la base d'un lien métaphorique avec une autre lexie du même vocable. En voici deux exemples :

- (53) La rencontre est un vrai désastre et Mulan est découragée. Ne le seriez-vous pas si on vous disait être une *disgrâce*? [sujet G-10]
 (54) Il est étrange que des gens qui ont tous *transités* (sic) par le secondaire ne se souviennent que partiellement de la réalité. [sujet A-18]

Dans la phrase (53), l'auteur a construit un usage métaphorique de DISGRÂCE sur une acception du vocable signifiant 'événement malheureux'. Le sujet G-10 a en quelque sorte « étiré » ce sens de sorte qu'il puisse s'appliquer à une personne qui serait la source de nombreux problèmes (ou événements malheureux). Dans (54), le verbe TRANSITER est utilisé pour signifier 'passer' et donne l'impression que l'auteur perçoit le secondaire comme un lieu de transition dans le cheminement scolaire. Cette utilisation de TRANSITER ne semble pas complètement saugrenue ; à la limite, nous pourrions croire que la personne a préféré le verbe TRANSITER au verbe PASSER, qui paraît moins recherché.

La phrase (55) contient un cas clair de métonymie :

- (55) ...des anneaux et des barres trapèzes pour la gymnastique quotidienne ; l'indien et le pirate servant de structure à d'autres balançoires, un *ballon-panier* et j'en passe. [sujet H-24]

Dans (55), le sujet a recours à la lexie BALLON-PANIER pour désigner le panier utilisé dans ce sport. Il désigne la partie (la pièce d'équipement) par le tout (le nom du sport), ce qui correspond à un cas de métonymie. Nous pouvons peut-être supposer que la personne tentait d'éviter d'écrire *un panier de ballon-panier*.

Finalement, le cas le plus intéressant de création d'une nouvelle unité lexicale est celui présent dans (56) :

(56) Personne ne recevra de coups de pieds *par faute d'être* plus marginal. [sujet K-02]

Dans ce cas-ci, l'auteur a créé la locution PAR FAUTE (de) à partir de l'acception de FAUTE signifiant 'manquement à une norme'. Nous pouvons supposer qu'il y a eu une certaine contamination de FAUTE (de). Pourtant, les sens des deux lexies n'ont rien en commun, FAUTE (de) signifiant 'par manque de' et PAR FAUTE (de) signifiant 'sa seule faute, son seul défaut étant de'.

Bizarrement, aucun des cas de création d'une nouvelle acception d'un vocable relevés dans le corpus ne correspond au cas de figure qui nous avait poussé à introduire cette classe d'erreurs : l'ajout d'un actant sémantique (voir II-2.3.4).

2.5.3 LES CAS DE MAUVAISES DÉRIVATIONS MORPHOLOGIQUES

Un seul parmi tous les problèmes lexicaux relevés dans le corpus a été considéré comme un cas « pur » de mauvaise dérivation morphologique :

(57) Cette cabane, on pouvait l'*emplacer* n'importe où (sic)... [sujet H-07]

Dans (57), le verbe *EMPLACER est clairement dérivé du nom EMPLACEMENT. Il est un peu étrange que la personne ait effectué la dérivation morphologique dans ce sens, puisque c'est généralement le nom qui est dérivé du verbe. Dans les faits, EMPLACEMENT est plutôt le nom typique utilisé pour exprimer le troisième actant sémantique du verbe PLACER :

X place Y sur/à Z, donc Z = emplacement

Nous verrons à la section consacrée aux cas de double codage que les deux autres cas de mauvaise dérivation morphologique recueillis ont aussi été perçus comme des cas de création d'une nouvelle acception d'un vocable.

2.6 LES CAS DE DOUBLE CODAGE

Notre analyse a révélé huit cas de double codage. En observant ces quelques cas, nous constatons que différentes raisons peuvent justifier un double codage. Par exemple, la phrase (58) a fait l'objet d'un double codage simplement parce que l'utilisation de *être d'accord que* entraîne ici deux problèmes.

(58) Tout d'abord, je **suis** parfaitement **d'accord qu'**au secondaire la discipline **soit** très présente... [sujet A-04]

Tout d'abord, il est clair qu'il y a un problème qui a trait au signifié, puisque le sens véhiculé par l'expression *être d'accord (que)* n'est pas celui que l'auteur voulait exprimer; c'est du moins ce qu'on peut déduire d'après le contexte. Il semble qu'il y ait eu confusion entre les expressions *être d'accord (que)* et *être d'accord (pour que)*, le sens de la deuxième semblant mieux correspondre au contexte. Comme la locution a en quelque sorte été déformée, nous avons codé ce problème comme un « non-respect du caractère figé d'une locution », mais nous aurions aussi pu parler d'impropriété, puisque le sens véhiculé par l'unité lexicale n'était pas le sens visé. Par ailleurs, l'emploi du subjonctif après la locution *être d'accord (que)* constitue un deuxième problème, qui correspond à notre classe « non-respect du mode verbal imposé par l'unité lexicale ».

Certains cas ont plutôt fait l'objet d'un double codage parce que deux interprétations semblaient pouvoir en rendre compte. C'est le cas du verbe FAIRE dans (59).

(59) Dans certaines écoles, on exige le port de l'uniforme [...] Je trouve cela très bien car cela **fait** moins de « taxage » et moins de préjugés au niveau du style vestimentaire... [sujet A-05]

Nous avons tout d'abord codé ce problème comme un cas d'imprécision. En effet, FAIRE constitue ici un choix de verbe peu recherché et le remplacer par un verbe comme ENTRAÎNER OU OCCASIONNER aurait rehaussé la phrase (59). Par contre, notre contre-codeuse nous a fait remarquer que le verbe FAIRE dans ce sens est fréquemment utilisé dans le langage parlé ; par conséquent, elle considérait qu'il s'agissait plutôt d'un problème de registre de langue. Comme les deux points de vue sont tous les deux valables, nous avons décidé de garder les deux codes. C'est le même cas pour le syntagme *Écoutez ce film* [sujet G-34], où ÉCOUTER peut être perçu soit comme un mauvais collocatif pour FILM, soit comme une unité lexicale appartenant à un niveau de langage familier en cooccurrence avec FILM, ou encore comme étant incompatible

sémantiquement avec *FILM* par rapport au typage de son deuxième actant syntaxique, qui devrait dénoter quelque chose qui est écouté (plutôt que regardé). Il peut donc arriver qu'il y ait triple codage!

Certains types de problèmes lexicaux semblent pouvoir être codés doublement dans la majorité des cas. Par exemple, les anglicismes sont très souvent perçus comme des impropriétés; nous discuterons plus bas de la possibilité de considérer les anglicismes comme une sous-classe d'impropriétés (voir IV-2.9.1). Les cas de mauvaise dérivation morphologique pourront aussi être perçus comme des créations d'une nouvelle acception d'un vocable chaque fois que la forme produite correspondra à une forme déjà existante. Par exemple, c'est le cas de *blanchissage d'argent* [sujet G-32] au sens de 'blanchiment d'argent' qui a été dérivé du verbe *BLANCHIR* (de l'argent), mais dont la forme correspond au vocable déjà existant *BLANCHISSAGE*.

Comme nous l'avions mentionné dans notre méthodologie, nous ne croyons pas que le double codage représente un problème en soi ou qu'il remette en doute nos classes d'erreurs. D'ailleurs, il est probable qu'un contre-codage sur le corpus complet aurait entraîné davantage de cas de double codage, puisqu'il aurait mis en lumière de nouvelles perspectives sur certains problèmes. Nous pensons par contre que certains cas de double codage indiquent qu'il y a plusieurs plans ou niveaux d'analyse des problèmes lexicaux. Il est très difficile de toujours s'en tenir exclusivement à une description structurale de l'erreur.

(60) *Je me rappellerais* (sic) toujours *de cet endroit* et de tout ce qu'on y a fait. [sujet H-27]

Par exemple, d'un point de vue structural, il faudrait considérer que l'utilisation transitive directe de *RAPPELER* dans (60) constitue un problème de non-respect du régime, puisque qu'il s'agit d'un verbe transitif direct. Par contre, il est évident que le sujet H-27 a utilisé cette construction du verbe *RAPPELER* car il l'a souvent entendue (ou même lue) ; on pourrait donc parler d'un problème de registre de langue. Mais cette deuxième interprétation ne relève-t-elle de l'explicatif plutôt que du structural? C'est pourquoi nous croyons qu'il existe différents plans d'analyse des problèmes lexicaux et qu'il est difficile de faire abstraction du niveau explicatif lorsqu'on s'intéresse au niveau descriptif.

2.7 LES ERREURS D'INATTENTION

En cours de codage, nous avons cru bon ajouter une classe « erreur d'inattention ». Comme l'ensemble des textes de notre corpus ont été rédigés par des étudiants francophones, certaines erreurs nous paraissaient exclues d'emblée. De plus, la majorité des textes étant écrits à l'ordinateur, nous pouvions facilement supposer que certaines erreurs étaient en fait des fautes de frappe. Considérer ces cas comme de réelles erreurs témoignant de lacunes relatives au lexique aurait selon nous faussé notre analyse. Les phrases (61) et (62) contiennent deux des cinq erreurs d'inattention relevées dans le corpus.

(61) ... les jeunes, pour la plupart, sont **portés ne pas faire** leurs leçons et devoirs...
[sujet A-05]

(62) Ensuite, nous escaladions la pente afin de nous **rentre** au sommet de la montagne.
[sujet H-03]

2.8 LE CAS LITIGIEUX

Des 391 problèmes lexicaux relevés dans notre corpus, un seul semblait ne pas correspondre aux types d'erreurs que nous avions anticipées. Voici l'extrait qui le contenait :

(63) ... les promenades en chaloupe à la poursuite de la **queue de l'arc-en-ciel**... [sujet H-09]

En effet, il semble que la lexie QUEUE ne soit pas vraiment appropriée pour désigner la partie d'un arc-en-ciel qui touche le sol; il serait plus juste de parler du *pied de l'arc-en-ciel*. En effectuant une recherche sur le Web, nous retrouvons quatre occurrences de « la queue de l'arc-en-ciel » et 88 occurrences de « le pied de l'arc-en-ciel ». De la même façon que l'unité lexicale QUEUE sert à désigner l'« extrémité » d'une comète, la lexie PIED semble être la façon standard de nommer l'« extrémité » d'un arc-en-ciel. Il s'agit d'un phénomène lexical qui s'apparente à la collocation en ce sens que le choix d'une lexie pour exprimer un certain sens est conditionné par le choix préalable d'une autre lexie. Dans ce cas-ci, par contre, la cooccurrence des deux lexies n'est pas nécessaire et elles appartiennent plutôt à un même champ lexical. La lexicologie explicative et combinatoire tient aussi compte de ce type de relations lexicales paradigmatiques, qu'elle appelle « dérivations sémantiques ». Les dérivations

sémantiques, aussi encodées par des fonctions lexicales comme les collocations, permettent d'associer des lexies qui présentent un lien sémantique (par exemple des synonymes ou des antonymes) ou encore un lien conceptuel (par exemple LACET et CHAUSSURE). Nous n'avions pas cru bon introduire une classe d'erreurs de dérivation sémantique dans notre typologie, car nous ne croyions pas trouver de tels cas dans les textes. Par contre, la phrase (63) semble prouver le contraire et nous pensons qu'il serait pertinent d'ajouter une classe « emploi d'un mauvais dérivé sémantique » à notre typologie. Ainsi, nous pourrions éliminer le seul cas que ne n'avons pas pu coder lors de notre première analyse.

2.9 SUGGESTIONS RELATIVES À LA TYPOLOGIE DES PROBLÈMES LEXICAUX

Le fait d'avoir utilisé notre typologie pour coder près de 400 problèmes lexicaux nous a bien sûr permis de constater certains problèmes et d'envisager certaines modifications à y apporter. Voici donc une version révisée de notre grille, qui selon nous permettrait une caractérisation plus fine et précise des problèmes lexicaux. Les sections suivantes expliqueront les changements suggérés.

Tableau IX. Typologie structurale des problèmes lexicaux (révisée)

TYPE	CAS		
A. Altération d'une des dimensions de l'unité lexicale	1. <i>Forme</i>	Erreur d'orthographe lexicale	
		Barbarisme	
	2. <i>Sens</i>	Non-respect d'une composante de sens	Non-respect du typage des actants.
			Non-respect d'une différence spécifique
	3. <i>Propriétés de combinatoire grammaticale</i>	Régime	Non-respect de la structure actancielle de la lexie
			Utilisation d'une mauvaise préposition
			Non-respect du mode verbal imposé par l'unité lexicale
		Autres propriétés grammati- cales inhérentes à la lexie	Non-respect du genre nominal
			Non-respect de la partie du discours
			Non-respect du caractère invariable
			Utilisation d'un mauvais verbe auxiliaire
		Non-respect d'une autre propriété grammaticale	Non-respect d'une autre propriété grammaticale
			Déformation de la locution
		4. <i>Propriétés de combinatoire lexicale</i>	Problème de collocation
	Choix d'un mauvais collocatif (collocation inexistante)		
	Choix d'un collocatif valable, mais dont le sens véhiculé n'est pas le bon		
	Non-respect des contraintes d'emploi du collocatif		
Choix d'un collocatif qui appartient à une autre lexie du même vocable			
Emploi d'un mauvais dérivé sémantique			
B. Utilisation d'une lexie fictive	1. Création d'une nouvelle acception d'un vocable		
	2. Impropropriété	Accidentelle	
		Anglicisme	
		Paronyme	
		Classe résiduelle	
3. Mauvaise dérivation morphologique	Création d'un signifiant fictif		
	Création d'une nouvelle acception d'un vocable		
C. Usage douteux en fonction du contexte	1. Changement de registre de langue		
	2. Mélange de codes (variantes dialectales)		
	3. Répétition : excessive ~ en écho		
	4. Imprécision : périphrase ~ mot générique		
	5. Pléonasme		

2.9.1 LES PROBLÈMES D'ALTÉRATION DU SENS ET DES PROPRIÉTÉS DE COMBINATOIRE GRAMMATICALE

Cette section de la typologie n'a pas subi de grands changements en ce qui concerne le contenu, mais certaines modifications ayant trait à la structure devraient en améliorer la cohérence.

Premièrement, comme nous l'avons mentionné à la section IV-1.1.1, nous avons réorganisé les problèmes relatifs à une altération du signifié. Les deux types de transgressions possibles, « non-respect du typage des actants » et « non-respect d'une différence spécifique », sont maintenant considérés comme des transgressions d'une composante de sens.

Nous avons aussi regroupé dans la classe « régime » les trois types d'erreurs suivants : « non-respect de la structure actancielle de la lexie », « utilisation d'une mauvaise préposition » et « non-respect du mode verbal imposé par l'unité lexicale ». La classe « non-respect d'une propriété inhérente à l'unité lexicale » a été rebaptisée « non-respect des autres propriétés de l'unité lexicale », afin de rendre compte du fait que le régime constitue aussi une propriété de l'unité lexicale. Elle gagnerait aussi à être divisée en sous-classes correspondant aux types d'erreurs trouvées : invariabilité, genre nominal et partie du discours. Une classe « résiduelle » pourrait regrouper les cas pour lesquels la propriété transgressée est plus spécifique ou particulière. Les problèmes de « non-respect » du caractère figé d'une locution pourraient aussi être divisés en deux groupes : ceux présentant une « déformation de la locution » et ceux découlant d'une « manipulation grammaticale incompatible avec la locution » (voir IV-2.2.3). De telles divisions permettraient une catégorisation plus fine des problèmes relatifs aux propriétés de combinatoire grammaticales.

2.9.2 LES CAS D'UTILISATION D'UNE LEXIE FICTIVE

La section B de notre typologie gagnerait à subir certains changements de division en ce qui concerne les classes « impropiété » et « mauvaise dérivation morphologique ». Nous avons évoqué au point IV-2.5.1 la possibilité d'introduire certaines sous-classes dans les impropiétés, puisque nous avons constaté que les impropiétés relevées n'étaient pas toutes de même nature.

Par conséquent, il y aurait peut-être lieu d'effectuer une subdivision entre les impropriétés « pures » ou « accidentelles » – par exemple, EXPRESSIF pour signifier 'turbulent' –, celles qui sont phonétiquement motivées (les paronymes) – par exemple INVASION pour signifier 'évasion' – et celles qui ont atterri là faute de diagnostic clair – par exemple PRÉCIS dans *un élève élevé de façon stricte et ²précise*. Les problèmes du dernier type seraient classés dans une classe résiduelle qui contiendrait les problèmes de sens pour lesquels il est difficile de cerner la composante de sens transgressée. Nous avons aussi mentionné, au point IV-1.6, qu'il serait possible de considérer les anglicismes comme des impropriétés. Par exemple, l'anglicisme ACADEMIQUE au sens de 'scolaire' constitue une impropriété, puisque que la forme *académique* ne correspond pas au sens 'scolaire' en français. Par contre, tous les anglicismes ne constituent pas des impropriétés. Nous n'avons qu'à penser aux anglicismes relatifs au régime, comme l'emploi de la préposition AVEC auprès de l'adjectif SATISFAIT dans *Il est satisfait avec son emploi* (*He is satisfied with his job.*). Il serait donc pertinent d'ajouter une sous-classe « anglicisme » dans les impropriétés, mais nous devons quand même conserver notre classe « mélange de codes » pour catégoriser les autres types d'anglicismes. Comme nous avons constaté que les mauvaises dérivations morphologiques pouvaient aussi être codées « création d'une nouvelle acception d'un vocable » lorsque la forme produite correspond à un signifiant déjà existant, nous proposons d'introduire deux sous-classes dans la classe « mauvaise dérivation morphologique ». Les deux classes suggérées permettront d'établir une distinction entre les cas où un signifiant inexistant est généré (par exemple EMPLACER [sujet -07]) et ceux pour lesquels la forme produite correspond à un vocable existant (comme BLANCHISSAGE (d'argent) [sujet G-32]).

2.9.3 LES PROPRIÉTÉS DE COMBINATOIRE LEXICALE

En ce qui concerne la partie « propriétés de combinatoire lexicale » de notre typologie, il y aurait lieu d'ajouter certaines classes à la seule déjà existante : « utilisation d'un mauvais collocatif ». En effet, lors du codage des problèmes de collocations, nous nous sommes aperçu qu'il en existe en fait plusieurs types. Nous avions à l'origine prévu un seul cas (« choix d'un mauvais collocatif »), celui où le collocatif employé ne devrait pas apparaître en cooccurrence avec la base choisie. L'exemple suivant est un cas d'erreur de collocation de ce type :

(64) Elle doit également travailler à son évolution tout en **appliquant un** plus grand **respect** de soi et d'autres personnes. [sujet K-01]

Nous avons par contre rencontré des problèmes de collocation différents du cas ci-haut. Par exemple, dans (65), nous avons une collocation existante (*avoir de la discipline*), mais dont le sens ne correspond pas à celui que l'auteur semblait vouloir exprimé (« choix d'un collocatif valable, mais dont le sens véhiculé n'est pas le bon »). En effet, pour exprimer le « bon » sens ici, l'auteur aurait dû employer un collocatif comme IMPOSER.

(65) ... les parents devraient **avoir beaucoup plus de discipline** à la maison avec les jeunes. [K-04]

Un troisième cas d'erreur de collocation que nous avons rencontré est celui où la collocation est tout à fait correcte, mais où certaines contraintes imposées par le collocatif ne sont pas respectées. Dans (66), le choix du collocatif PORTER est excellent et correspond à une collocation typique de la base ATTENTION. Par contre, l'emploi du déterminant *une* semble poser problème ici. En effet, avec le collocatif PORTER, le complément direct (la base) est introduit sans déterminant (*porter attention à quelque chose*), à moins bien sûr qu'un adjectif ne vienne modifier la base (*porter une attention particulière à quelque chose*).

(66) Ici, je voudrais **porter une attention** à la transparence du personnel quant au respect des mêmes règles. [sujet K-09]

Il existe aussi un dernier type de problème de collocation que nous n'avons pas rencontré dans notre corpus, mais dont nous avons entendu un exemple dans un cours. Dans ce cas-ci, le collocatif choisi correspond à une collocation d'une autre unité lexicale du même vocable³⁴. Dans (67), la personne emploie l'acception de FAMILLE qui signifie 'groupe d'éléments appartenant au même ensemble' (Antidote Prisme, 2004 : article FAMILLE), alors que la collocation qu'elle utilise correspond plutôt à l'acception 'ensemble de personnes vivant quotidiennement les unes avec les autres dans le même domicile et [Spécialement], le père, la mère et les enfants' (Antidote Prisme, 2004 :

³⁴ Ce cas de figure correspond d'ailleurs à l'un des types de « greffes collocationnelles » proposés par Polguère (à paraître). Les *greffes collocationnelles* sont « *grosso modo* [des] collocations où le collocatif semble « emprunté » à une autre collocation – généralement, une collocation dont la base est sémantiquement proche de la base à laquelle le collocatif emprunté est greffé. » (Polguère, à paraître: 2) Un des exemples fournis est *chaleur de plomb*.

article FAMILLE) qui est nettement plus courante. Avec l'acception utilisée, le collocatif CRÉER aurait été plus approprié!

(67) Vous allez dans le menu « outils » pour *fonder une famille* de codes.

Comme nous l'avons mentionné au point IV-2.8, il pourrait aussi être important d'introduire dans cette zone de la typologie une classe « utilisation d'un mauvais dérivé sémantique ».

CONCLUSION

Cette analyse des résultats nous a permis de mettre en lumière certaines informations intéressantes par rapport aux problèmes lexicaux présents dans les textes d'étudiants universitaires. Tout d'abord, le nombre élevé de problèmes relatifs aux propriétés de combinatoire grammaticale devrait nous convaincre de la nécessité d'accorder plus d'importance à cette dimension des unités lexicales dans l'enseignement du français langue première. Favoriser l'utilisation stratégique et efficace des ouvrages de référence pourrait diminuer de beaucoup le nombre d'erreurs de ce type, ainsi que la présence d'impropriétés, qui constituent aussi un type majeur de problème lexical.

Le fait d'introduire un deuxième axe d'analyse nous a aussi permis de constater qu'il est généralement possible de trancher sur la question du statut erroné ou non d'un problème lexical. Les seuls cas pour lesquels il était plus difficile de parler d'erreurs étaient, comme nous l'avions prévu, ceux relevant de normes autres que la norme prescriptive (cf. usage douteux en fonction du contexte). La totalité des problèmes relatifs à une altération du signifiant (barbarismes et erreurs d'orthographe lexicale), des problèmes d'inattention, de ceux découlant du non-respect du mode verbal imposé par l'unité lexicale et des lexies fictives créées par dérivation morphologique ont été considérés comme des erreurs. Il y a donc lieu de supposer que des problèmes relevant de ces classes ont de fortes chances d'être très souvent perçus comme des erreurs.

Le fait d'avoir testé notre typologie sur un corpus de réels textes d'étudiants nous a permis de constater que les classes que nous avons déterminées d'un point de vue théorique étaient valables, mais requéraient, dans certains cas, quelques modifications. Par cet exercice, notre typologie a gagné en validité et en pertinence. La mettre à l'épreuve sur un corpus plus large, incluant cette fois des textes écrits par des étudiants n'ayant pas le français comme langue maternelle, permettrait sans doute de la raffiner

davantage. Probablement que l'analyse de problèmes lexicaux issus de textes d'élèves plus jeunes – comme ceux du primaire et du secondaire – nous amènerait aussi à proposer certains réaménagements dans notre grille, mais nous croyons que ces changements seraient mineurs et laisseraient intacte la structure de notre typologie.

Finalement, la réalisation de notre premier objectif de recherche nous aura aussi démontré qu'il est pratiquement impossible, d'un point de vue didactique, de se limiter à une analyse structurale des erreurs lexicales, puisque le simple exercice de caractérisation de problèmes de ce type fait souvent émerger des scénarios d'explication ou de diagnostic de leur(s) source(s). En effet, nous avons démontré que l'identification d'un problème lexical à une classe particulière implique parfois une certaine prise de position par rapport à la cause possible du problème (voir IV-2.6). Par exemple, un anglicisme relatif au régime comme l'emploi d'une mauvaise préposition n'est, d'un point de vue purement structural, qu'un problème de préposition régie; en classant ce problème parmi les anglicismes, nous supposons que le choix de cette préposition particulière a été fait sur la base d'une analogie avec une expression de l'anglais. Notre typologie des problèmes lexicaux est donc structurale dans la mesure où elle a été élaborée dans une perspective descriptive, mais son utilisation pour caractériser des problèmes lexicaux n'est pas exempte de considérations explicatives.

**V. LES STRATÉGIES DE
RÉSOLUTION DE PROBLÈMES
LEXICAUX DES ÉTUDIANTS
UNIVERSITAIRES**

1. PRÉSENTATION DES RÉSULTATS (OBJECTIF 2)

Notre deuxième objectif de recherche est le suivant :

Décrire les stratégies de résolution de problèmes lexicaux d'étudiants universitaires, notamment en fonction des types de problèmes rencontrés.

Dans le but de décrire ces stratégies, nous avons analysé huit entrevues menées auprès d'étudiants universitaires. Comme nous l'avons mentionné dans notre méthodologie (voir III-3.5), l'analyse des entrevues a été effectuée à deux niveaux : les séquences et les commentaires. Par conséquent, nous présenterons nos résultats en deux parties.

1.1 LES SÉQUENCES : RÉSULTATS

L'analyse des entrevues a révélé un total de 321 séquences, c'est-à-dire que les sujets se sont intéressés à 321 problèmes. Comme le test soumis aux sujets comportait deux parties correspondant à des tâches différentes, nous analyserons les séquences en deux temps. La section V-1.1.1 s'intéressa exclusivement à celles du premier volet du test et déterminera quels problèmes ont été les plus repérés. La section V-1.1.2 portera sur l'ensemble des séquences et concernera la résolution des problèmes discutés, qu'ils appartiennent à l'une ou l'autre des parties du test.

1.1.1 PARTIE 1 : REPÉRAGE D'ERREURS LEXICALES

La première partie du test comportait 17 problèmes lexicaux insérés dans un texte d'une page. Les sujets avaient pour tâche de lire le texte, de repérer les problèmes et d'en discuter. Comme aucune indication n'avait été donnée quant au fait que la recherche s'intéressait exclusivement aux erreurs lexicales, certains sujets ont pu s'attarder sur des problèmes relevant d'autres dimensions (syntaxe, cohérence textuelle, orthographe grammaticale, etc.). Il est par contre important de rappeler que seuls des problèmes lexicaux étaient présents dans ce texte.

La première partie du test a été à l'origine de 201 séquences, soit 62,6 % du total des problèmes abordés dans les deux tâches. De ces 201 séquences, 65 portaient sur des problèmes non lexicaux, ce qui représente près du tiers (32,3 %). Nos transcriptions du

premier volet du test comportent donc 136 séquences s'intéressant à des problèmes lexicaux; 80 correspondent à des problèmes cibles³⁵, alors que 56 concernent des problèmes lexicaux non cibles, c'est-à-dire des éléments ne présentant pas d'erreur. Si les huit équipes avaient traité l'ensemble des 17 problèmes lexicaux présents dans le test, nous aurions obtenu 136 séquences. Comme seulement 80 séquences portaient sur ces problèmes cibles, nous pouvons conclure que 58,8 % des problèmes lexicaux du texte ont été repérés.

Le tableau X présente les résultats concernant le repérage des problèmes cibles de la première partie du test. La première colonne indique le type de problème, alors que la seconde montre l'extrait concerné. La troisième colonne mentionne le nombre (et le pourcentage) d'équipes qui ont relevé le problème en question.

Tableau X. *Problèmes cibles repérés dans la première partie du test*

Type de problème	Extrait	Repérés
Usage d'une mauvaise préposition	<i>sont plutôt portés de manquer</i>	8 100 %
Non-respect du caractère figé d'une locution	<i>trouver le juste et bon milieu</i>	8 100 %
Non-respect du genre nominal	<i>différents idoles</i>	7 87,5 %
Erreur d'orthographe lexicale	<i>d'avantage</i>	6 75,0 %
Mauvaise collocation	<i>employer leur autorité</i>	6 75,0 %
Impropriété	<i>la paresse attaque davantage</i>	6 75,0 %
Non-respect d'une différence spécifique	<i>Pour en terminer avec la discipline</i>	6 75,0 %
Registre de langue inapproprié	<i>déprimeraient moins de devoir</i>	6 75,0 %
Pléonasme	<i>plus les générations sont devenues davantage permissives</i>	6 75,0 %
Mauvaise collocation	<i>appliquer un plus grand respect</i>	5 62,5 %
Impropriété	<i>deux groupes corroborent (= 'collaborent')</i>	5 62,5 %
Non-respect du mode verbal imposé par l'unité lexicale	<i>bien que le fouet n'est plus utilisé</i>	5 62,5 %
Anglicisme	<i>l'opportunité d'apprendre</i>	2 25,0 %
Anglicisme	<i>mettre l'emphase</i>	2 25,0 %
Impropriété	<i>imposer des conséquences (= 'punitions')</i>	1 12,5 %
Usage d'une mauvaise préposition	<i>l'intérêt des jeunes envers leurs études</i>	1 12,5 %
Répétition	<i>en s'intéressant aux intérêts</i>	0 0 %

Les trois problèmes les plus repérés sont des problèmes relevant d'une transgression des propriétés de combinatoire grammaticale. Deux de ces problèmes ont été relevés

³⁵ Par « problèmes cibles », nous désignons les problèmes que nous voulions voir traités par les sujets. Ces problèmes constituaient les seules erreurs ou maladroites lexicales présentes dans le test.

par l'ensemble des équipes, alors qu'une seule n'a pas discuté le troisième. La seule erreur d'orthographe lexicale a été repérée par les trois quarts des équipes, et les deux problèmes de collocation ont été traités par respectivement 75 % et 62,5 % des dyades. Les quatre erreurs relatives au sens figurant dans le texte (3 impropriétés et 1 non-respect d'une différence spécifique) présentent de grandes différences quant à leur taux de repérage; deux ont été relevées par 75 % des équipes, une par 62,5 %, alors qu'une seule équipe (12,5 %) a repéré la dernière. Six dyades sur 8 (75 %) ont noté le cas de pléonasme et celui d'utilisation d'une unité lexicale appartenant au registre familier. Soixante-deux et cinq dixièmes pour cent (62,5 %) des sujets ont remarqué l'usage erroné du mode indicatif après la locution BIEN QUE. Les anglicismes ont été peu repérés : par 2 équipes sur 8 dans les deux cas. Des deux problèmes d'utilisation d'une mauvaise préposition, un a été relevé par tous les sujets, alors que seule une équipe (12,5 %) a mentionné l'autre.

Finalement, le cas de répétition n'a été pointé par personne. Nous tenterons de comprendre, au point V-2.1.1, pourquoi certains problèmes ont été moins repérés que d'autres.

1.1.2 LA RÉOLUTION DES PROBLÈMES LEXICAUX

Nous avons vu que 80 des 201 séquences de la première partie du test portent sur des problèmes cibles, soit 39,8 %. Parmi ces 80 problèmes, 66 ont été résolus correctement (82,5 %), alors que 14 n'ont pas été résolus (17,5 %). Si nous considérons l'ensemble des 136 problèmes lexicaux qui auraient pu être repérés et traités par les équipes (17 problèmes x 8 équipes), nous constatons que seulement 48,5 % des problèmes présents dans le texte ont été correctement résolus. Quarante-six séquences portant sur les problèmes cibles (57,5 %) présentaient un contenu jugé suffisant à la résolution du problème traité et 34 (42,5 %) ont été considérées insuffisantes.

Pour la deuxième partie du test, la situation est un peu différente, puisque chacune des quinze phrases indépendantes soumises aux sujets contenait un problème lexical en caractère gras. Par conséquent, tous les problèmes présents dans cette seconde partie ont été discutés; elle est donc à l'origine de 120 séquences qui correspondent aux 120 problèmes étudiés (15 problèmes x 8 équipes). La totalité de ces séquences portent sur des problèmes cibles. Parmi elles, 95 (79,2 %) ont conduit à une bonne résolution du problème et 25 correspondent à des problèmes mal résolus (20,8 %). Quatre-vingt-six

séquences de la deuxième partie du test ont été codées suffisantes (71,7 %), alors que 34 (28,3 %) ne véhiculent pas des stratégies suffisantes en regard du problème à résoudre.

Nous obtenons donc un total de 200 séquences relatives à des problèmes cibles pour l'ensemble du test. Des 200 problèmes traités, 161 ont été bien résolus (80,5 %), alors que 39 n'ont pas été résolus correctement (19,5 %). Le contenu de 132 de ces séquences (66,0 %) a été jugé suffisant à la résolution du problème traité et 68 séquences (34,0 %) présentaient des stratégies considérées insuffisantes.

Comme 80,5 % des problèmes cibles traités ont été bien résolus et que seulement 66,0 % des séquences y correspondant présentent des stratégies suffisantes, il est évident qu'une séquence insuffisante peut tout de même conduire à la correction du problème. En effet, parmi les 161 problèmes bien résolus, 34 correspondent à des séquences codées insuffisantes, ce qui représente plus du cinquième des cas (21,1 %). Les 127 autres cas bien résolus l'ont été à l'issue de séquences considérées suffisantes (78,9 %). De la même façon, 34 des problèmes mal résolus (87,2 %) correspondent à des séquences insuffisantes. Finalement, il existe cinq cas plus marginaux où le problème n'a pas été résolu en dépit du fait qu'il soit associé à une séquence présentant des stratégies suffisantes (12,8 %).

Le tableau XI présente des résultats concernant les 200 séquences portant sur des problèmes cibles. La première colonne indique le type de problème traité et la seconde mentionne le nombre de séquences s'intéressant à un problème du type en question; cette information nous permettra de relativiser les chiffres des colonnes suivantes. En effet, des résultats concernant un type de problème n'ayant fait l'objet que de 4 séquences ont certainement moins de poids que ceux découlant de l'analyse de plus de 10 séquences. Les colonnes *bien résolus* et *mal résolus* présentent le nombre de séquences, parmi celles concernant des problèmes du type correspondant, ayant conduit à de bonnes ou de mauvaises résolutions (avec leur pourcentage respectif). Finalement, les deux dernières colonnes indiquent le nombre de séquences présentant des stratégies suffisantes ou insuffisantes à la résolution du problème concerné.

Tableau XI. Problèmes cibles bien et mal résolus et statuts des séquences correspondantes

Type	Nb séqu.	Bien résolus		Mal résolus		Suffisantes		Non suffisantes	
Non-respect d'une propriété inhérente	15	15	100,0 %	0	0,0 %	14	93,3 %	1	6,7 %
Utilisation d'une mauvaise préposition	9	9	100,0 %	0	0,0 %	4	44,4 %	5	55,6 %
Périphrase	8	8	100,0 %	0	0,0 %	8	100,0 %	0	0,0 %
Non-respect du caractère figé d'une locution	16	15	93,8 %	1	6,3 %	12	75,0 %	4	25,0 %
Problème de forme	14	13	92,9 %	1	7,1 %	12	85,7 %	2	14,3 %
Registre de langue	30	25	83,3 %	5	16,7 %	23	76,7 %	7	23,3 %
Pléonasme	6	5	83,3 %	1	16,7 %	4	66,7 %	2	33,3 %
Problème de sens	50	40	80,0 %	10	20,0 %	39	78,0 %	11	22,0 %
Anglicisme	4	3	75,0 %	1	25,0 %	2	50,0 %	2	50,0 %
Mauvaise collocation	19	12	63,2 %	7	36,8 %	9	47,4 %	10	52,6 %
Non-respect du mode verbal	5	3	60,0 %	2	40,0 %	0	0,0 %	5	100,0 %
Non-respect de la structure actancielle	16	9	56,3 %	7	43,8 %	2	12,5 %	14	87,5 %
Utilisation d'un mauvais auxiliaire	8	4	50,0 %	4	50,0 %	3	37,5 %	5	62,5 %
TOTAL	200	161	80,5 %	39	19,5 %	132	66,0 %	68	34,0 %

Trois types de problèmes ont été résolus correctement chaque fois qu'ils ont été traités : ceux ayant trait à la transgression d'une propriété inhérente de l'unité lexicale (ici genre nominal et partie du discours), ceux présentant l'emploi d'une mauvaise préposition et le problème de périphrase. Si le nombre de séquences jugées suffisantes est élevé pour les problèmes relatifs aux propriétés inhérentes et à l'emploi d'une périphrase – respectivement 93,3 % et 100 % –, il n'est que de 44,4 % dans les cas qui concernent l'usage de mauvaises prépositions.

Les problèmes découlant du non-respect du caractère figé d'une locution ou de l'emploi d'unités lexicales appartenant à un registre de langue familier, les problèmes de forme et le pléonasme ont été bien résolus dans la grande majorité des cas où ils ont été traités (de 83,3 % à 93,8 %). Les séquences suffisantes concernant ces quatre types de problèmes présentent des pourcentages moins élevés, mais sont somme toute assez proches des taux de résolution (inférieurs de 6,6 % à 18,8 %).

Les problèmes de sens sont à l'origine du plus grand nombre de séquences et ont été bien résolus dans 80,0 % des cas et présentent des stratégies suffisantes dans 78,0 % des cas.

Les anglicismes ou les mauvaises collocations ont été bien corrigés dans 75,0 % et 63,2 % des cas. Par contre, les séquences concernant ces problèmes n'ont été

considérées suffisantes qu'environ la moitié du temps (50,0 % et 47,4 %). Il faut cependant relativiser ces résultats par rapport aux anglicismes, puisqu'ils découlent de l'analyse d'un nombre très restreint de séquences (4).

Le cas de non-respect du mode verbal imposé par l'unité lexicale BIEN QUE attire l'attention. Si 60,0 % des équipes qui l'ont traité l'ont bien résolu, aucune n'a su déployer des stratégies suffisantes pour arriver à proposer une solution correcte. Les cas de non-respect de la structure actancielle d'une unité lexicale présentent des résultats similaires : ils ont été bien résolus dans 56,3 % des cas, alors que seulement 12,5 % des séquences les concernant ont été jugées suffisantes.

Finalement, le cas d'utilisation d'un mauvais auxiliaire a été bien résolu par la moitié des équipes, mais la correction n'a été justifiée de façon satisfaisante que dans 37,5 % de cas.

1.2 LES COMMENTAIRES : RÉSULTATS

Lors de l'analyse de nos huit entrevues, 1531 commentaires ont été retenus, ce qui représente une moyenne de 4,8 commentaires par séquence. Nous rappelons que plusieurs verbalisations n'ont pas été codées parce qu'elles ne s'avéraient pas pertinentes en regard de notre analyse, qu'elles étaient incomplètes ou qu'elles reprenaient des éléments déjà présents dans d'autres verbalisations appartenant à la même séquence (voir III-3.4.2). Les commentaires seront analysés tous ensemble, sans distinction entre les deux parties du test, ni entre ceux portant sur des problèmes cibles ou non cibles. La section V-1.2.1 présentera un portrait général des commentaires, alors que la section V-1.2.2 portera plus spécifiquement sur le contenu de ces commentaires, en d'autres mots sur ce qui constitue véritablement les stratégies de résolution de problèmes lexicaux. Finalement, la section V-1.2.3 présentera les principales stratégies utilisées pour chaque type de problème lexical traité.

1.2.1 LES COMMENTAIRES : PORTRAIT GÉNÉRAL

Sur les 1531 commentaires, la très grande majorité étaient spontanés (1355, soit 88,5 %), c'est-à-dire qu'ils ont été émis naturellement par un sujet ou en réaction à une remarque de son coéquipier. Ce sont donc seulement 176 commentaires (11,5 %) qui ont été suscités par une intervention de l'intervieweur; ces interventions visaient un approfondissement d'un point déjà soulevé par les sujets dans 164 cas (10,7 %) et constituaient des pistes d'analyse du problème traité dans seulement 12 cas (0,8 %).

Le contenu de 962 verbalisations (62,8 %) a été jugé approprié pour la résolution du problème, alors que 329 (21,5 %) commentaires véhiculaient des informations non pertinentes ou erronées. Deux-cent-quarante commentaires (15,7 %) ont reçu le statut *indéfini* pour ce critère.

L'analyse a été l'étape de la résolution du problème à l'origine de près de la moitié des verbalisations (761, soit 49,7 %), alors que l'étape de proposition et d'analyse de solutions aux problèmes traités a entraîné 414 commentaires (27,0 %). Deux-cent-trente-neuf commentaires véhiculaient une prise de décision (15,6 %) et 117 concernaient le repérage (7,6 %). Le tableau suivant indique la répartition entre les commentaires appropriés, inappropriés et indéfinis en fonction de l'étape de résolution.

Tableau XII. Statuts des commentaires en fonction de l'étape de résolution du problème

Étape de la résolution	Total	Commentaires appropriés		Commentaires inappropriés		Commentaires indéfinis	
Repérage	117	49	41,9 %	12	10,3 %	56	47,9 %
Analyse	761	468	61,5 %	165	21,7 %	128	16,8 %
Proposition et analyse de solutions	414	268	64,7 %	110	26,6 %	36	8,7 %
Décision	239	177	74,1 %	42	17,6 %	20	8,4 %

En ce qui concerne l'emploi du métalangage dans les verbalisations, il était assez peu répandu : 190 (12,4 %) des commentaires en contenaient alors que 1341 (87,6 %) en étaient exempts. Un examen plus approfondi des commentaires contenant des éléments de métalangage a révélé l'emploi de 67 termes différents utilisés 248 fois. Nous verrons quels sont les éléments du métalangage utilisés lors des entrevues à la section V-2.2.1.

Le tableau XIII synthétise les résultats présentés plus haut.

Tableau XIII. *Données concernant les 1531 commentaires codés*

Critère	Valeurs possibles	Nombres de cas	% des 1531 commentaires
Spontanéité du commentaire	Spontané	1355	88,5 %
	Suscité – approfondissement	164	10,7 %
	Suscité – piste d’analyse	12	0,8 %
Étape de la résolution du problème à laquelle appartient le commentaire	Repérage	117	7,6 %
	Analyse	761	49,7 %
	Proposition et analyse de solution(s)	414	27,0 %
	Décision	239	15,6 %
Caractère approprié du commentaire	Approprié	962	62,8 %
	Inapproprié	329	21,5 %
	Indéfini	240	15,7 %
Présence de métalangage	Oui	190	12,4 %
	Non	1341	87,6 %

1.2.2 LES TYPES DE COMMENTAIRES

La section précédente nous a permis d'avoir un portrait global des commentaires émis par les sujets, mais pour véritablement décrire les stratégies de résolution des problèmes lexicaux, il faudra analyser plus en détails le contenu des verbalisations. En effet, en ayant une meilleure idée des types de commentaires constituant les séquences, nous pourrons mieux comprendre les stratégies mises en œuvre par les étudiants.

C'est le cinquième axe d'analyse des commentaires de notre grille de codage (voir tableau V, III-3.4.2) – *types de commentaires* – qui nous a permis d'en décrire le contenu. Des 1531 commentaires codés, seulement 161 (10,5 %) n'ont reçu aucun code concernant cet axe d'analyse; ces verbalisations concernent, pour la plupart, les étapes de repérage des problèmes et de prise de décision. Tous les autres commentaires (1370, soit 89,5 %) se sont vu attribuer soit au moins un code *type de commentaires*. Nous rappelons ce que nous avons mentionné dans notre méthodologie (voir III-3.4.2) : certains commentaires véhiculent à la fois une stratégie visant à résoudre le problème, par exemple le recours à un ouvrage de référence, et des éléments diagnostiques ou périphériques, par exemple un jugement global (« Corroborer : confirmer, appuyer ». Bon, c'correct. [équipe 1, séq. 12]). Les commentaires ayant reçu des codes de la classe *commentaire visant la résolution du problème* (1104, soit 72,1 % du total) sont beaucoup plus fréquents que les commentaires diagnostiques ou périphériques (375, soit 24,5 % du total). Parmi ces commentaires, 109 ont les deux types de codes, ce qui représente 7,1 % du total.

Le tableau suivant présente le nombre de commentaires de chaque type.

Tableau XIV. Types de commentaires : nombre et pourcentage des 1531 commentaires codés

Type de commentaire		Nombre de commentaires			% du total des commentaires
		conduisant à des problèmes bien résolus	conduisant à des problèmes mal résolus	total	
Visant la résolution du problème	proposition faisant partie d'un raisonnement (justifiée)	235	45	280	18,3 %
	consultation d'un ouvrage de référence	103	34	137	8,9 %
	recours au sens	113	21	134	8,8 %
	analyse grammaticale ou sémantique de la phrase	95	20	115	7,5 %
	recours au jugement du locuteur natif	50	21	71	4,6 %
	proposition catégorique (non justifiée)	46	13	59	3,9 %
	référence à des propriétés de combinatoire grammaticale	47	9	56	3,7 %
	comparaison de deux solutions (oral ou écrit)	49	7	56	3,7 %
	référence à la norme fonctionnelle	28	13	41	2,7 %
	référence à la forme du mot-forme	18	6	24	1,6 %
	énonciation d'une « règle »	16	5	21	1,4 %
	relecture	13	6	19	1,2 %
	illustration d'une bonne utilisation de l'unité lexicale traitée par un exemple	12	1	13	0,8 %
	référence à une autorité	11	1	12	0,8 %
	comparaison/analogie avec un autre problème	10	0	10	0,7 %
	mise en parallèle d'un bon et d'un mauvais exemple	7	3	10	0,7 %
	référence à une expression	9	0	9	0,6 %
	autre	7	1	8	0,5 %
	évitement	7	0	7	0,5 %
	recours à la paraphrase, la reformulation	6	0	6	0,4 %
	établissement d'un lien morphologique avec un autre mot-forme	3	2	5	0,3 %
	remplacement par une autre unité lexicale	5	0	5	0,3 %
	établissement d'un lien avec une unité lexicale de l'anglais	2	1	3	0,2 %
	apport d'une évidence linguistique	2	1	3	0,2 %
Diagnostic	jugement global	187	48	235	15,3 %
	jugement esthétique	39	10	49	3,2 %
Périphérique	commentaire relatif à la tâche	14	4	18	1,2 %
	expression d'incertitude	43	30	73	4,8 %

Les commentaires les plus fréquents (18,3 %) sont ceux véhiculant une proposition de solution au problème traité dans le cadre d'un raisonnement. La référence à un ouvrage de référence et le recours au sens d'une unité lexicale constituent des stratégies fréquentes dans la résolution de problèmes lexicaux (8,9 % et 8,8 %). Les sujets ont fait appel à l'analyse grammaticale ou sémantique de la phrase traitée dans 7,5 % des cas et se sont fiés à leur jugement de locuteur dans 4,6 % des cas. Trois et neuf dixièmes pour cent (3,9 %) des verbalisations dénotent une correction catégorique du problème sans aucune explication ou justification. La référence à des propriétés de combinatoire grammaticale et la comparaison de deux solutions occupent le septième rang des stratégies de résolution de problèmes lexicaux les plus fréquentes et représentent 3,7 % des cas. Dans 2,7 % des commentaires, les sujets ont fait référence à la norme fonctionnelle et ils se sont intéressés au signifiant de l'unité lexicale dans 1,6 % des verbalisations. L'énonciation d'une « règle » relative au problème discuté constituait 1,4 % des commentaires et 1,2 % présentaient une relecture du segment de texte corrigé.

Les onze autres types de commentaire de notre grille de codage ne représentent chacun que moins de 1,0 % des commentaires (entre 0,8 % et 0,2 %). Finalement, les verbalisations ne correspondant pas aux classes de notre grille de codage ont été classées dans la catégorie résiduelle *autre* et constituent 0,5 % des cas (8 commentaires).

Pour ce qui est des commentaires diagnostiques, les plus fréquents étaient ceux véhiculant un jugement global sur le problème traité ou sur une solution proposée pour y remédier (15,3 %). Les sujets ont posé un jugement esthétique sur le problème discuté dans 3,2 % de leurs commentaires. Finalement, 4,8 % des verbalisations contenaient une manifestation explicite d'incertitude et 1,2 % portaient sur la tâche à accomplir.

1.2.3 LES STRATÉGIES EN FONCTION DES TYPES DE PROBLÈMES LEXICAUX

Notre deuxième objectif de recherche a pour but de décrire les stratégies de résolution mises en œuvre par les étudiants selon les types de problèmes traités. La section précédente nous a permis d'observer les types de commentaires émis par les sujets et de constater les stratégies privilégiées de façon générale par ces derniers. Nous allons

maintenant observer si les stratégies déployées varient en fonction du type de problème rencontré. Pour ce faire, nous avons regroupé les commentaires émis par types de problèmes traités (problèmes de sens, mauvaises collocations, etc.). Comme nous nous intéressons aux stratégies proprement dites, nous avons ensuite retranché tous les commentaires ne visant pas directement la résolution du problème³⁶. L'ensemble des commentaires restants appartenait donc à la classe des commentaires visant la résolution du problème. Pour chaque type de problèmes, nous avons donc pu observer les stratégies de résolution les plus fréquentes³⁷. Ces observations figurent dans le tableau XV (ci-dessous) et sont ordonnées des types de problèmes les mieux résolus à ceux les moins bien résolus (voir V-1.1.2).

Pour huit types de problèmes sur treize, la consultation d'un ouvrage de référence figure parmi les stratégies les plus fréquentes. L'analyse grammaticale ou sémantique de la phrase traitée et le recours au sens sont les stratégies privilégiées pour cinq des types de problèmes, suivis par la référence à des propriétés de combinatoire grammaticale et le recours au jugement du locuteur natif (quatre types de problèmes). Deux types de problèmes lexicaux comptent la proposition catégorique parmi les stratégies les plus souvent mises en œuvre dans leur résolution : les problèmes de collocation et l'usage d'une mauvaise préposition. Finalement, quatre stratégies ne sont dominantes que pour un type particulier de problème : la comparaison de deux solutions pour les problèmes de sens, la référence à la forme du mot-forme pour les problèmes de forme, la référence à la norme fonctionnelle dans le cas des problèmes de registres de langue et la référence à une expression pour les problèmes de non-respect du caractère figé d'une locution.

De plus, le tableau XV nous révèle que la proportion de stratégies dominantes varie énormément d'un type de problème à l'autre. Par exemple, les stratégies qui se démarquent pour les problèmes d'utilisation d'une mauvaise préposition totalisent 60,9 % du total des stratégies mises en œuvre pour ce type de problème. Pour le problème d'utilisation d'un mauvais verbe auxiliaire, les stratégies récurrentes ne

³⁶ Les 161 verbalisations n'ayant reçu aucun code du critère *type de commentaires* n'ont donc pas été retenues pour cette étape de l'analyse, tout comme les commentaires exclusivement diagnostiques ou périphériques.

³⁷ Comme nous l'avons mentionné à la section III-3.4.2 de notre méthodologie, les commentaires véhiculant une proposition de solution dans le cadre d'un raisonnement font partie du processus normal de résolution de problèmes et ne sont pas d'un grand intérêt pour nous. D'ailleurs, les séquences relatives à tous les types de problèmes contenaient un nombre assez important de commentaires de ce type. Pour cette raison, nous n'avons pas cru bon indiquer cette stratégie de résolution dans le tableau XV.

représentent que 26,2 % de l'ensemble des stratégies observées. Nous verrons dans la section V-2.2.3 les implications de ces constatations.

Tableau XV. Stratégies les plus fréquentes selon les types de problèmes lexicaux

Type de problème	Stratégies les plus fréquentes	% des stratégies
Non-respect d'une propriété inhérente	référence à des propriétés de combinatoire grammaticale	31,6 %
	consultation d'un ouvrage de référence	17,5 %
Utilisation d'une mauvaise préposition	proposition catégorique (non justifiée)	17,1 %
	recours au jugement du locuteur natif	14,6 %
	référence à des propriétés de combinatoire grammaticale	14,6 %
	consultation d'un ouvrage de référence	14,6 %
Périphrase	analyse grammaticale ou sémantique de la phrase	38,5 %
Non-respect du caractère figé d'une locution	référence à une expression	12,9 %
	recours au sens	12,9 %
Problème de forme	référence à la forme du mot-forme	23,3 %
	consultation d'un ouvrage de référence	23,3 %
Changement de registre de langue	référence à la norme fonctionnelle	19,3 %
	consultation d'un ouvrage de référence	8,6 %
	recours au jugement du locuteur natif	7,9 %
Pléonasme	analyse grammaticale ou sémantique de la phrase	31,3 %
	recours au sens	9,4 %
	recours au jugement du locuteur natif	9,4 %
Problème de sens	recours au sens	20,9 %
	consultation d'un ouvrage de référence	17,7 %
	analyse grammaticale ou sémantique de la phrase	9,4 %
	comparaison de deux solutions (oral ou écrit)	6,7 %
Anglicisme	consultation d'un ouvrage de référence	33,3 %
Mauvaise collocation	consultation d'un ouvrage de référence	13,9 %
	proposition catégorique (non justifiée)	12,7 %
	recours au sens	10,1 %
Non-respect du mode verbal	recours au jugement du locuteur natif	16,7 %
	analyse grammaticale ou sémantique de la phrase	16,7 %
Non-respect de la structure actancielle	consultation d'un ouvrage de référence	13,3 %
	référence à des propriétés de combinatoire grammaticale	9,6 %
	analyse grammaticale ou sémantique de la phrase	9,6 %
	référence à des propriétés de combinatoire grammaticale	14,3 %
Utilisation d'un mauvais auxiliaire	recours au sens	11,9 %

2. INTERPRÉTATION DES RÉSULTATS (OBJECTIF 2)

La section qui suit analysera le contenu des entrevues réalisées pour atteindre de notre deuxième objectif : la description des stratégies de résolution de problèmes lexicaux des étudiants universitaires. Comme pour la présentation des résultats, nous traiterons de façon indépendante les séquences (section V-2.1) et les commentaires (section V-2.2).

2.1 LES SÉQUENCES

Comme pour la présentation des résultats, nous aborderons l'interprétation des résultats relatifs aux séquences en nous intéressant aux types de problèmes qui ont retenu l'attention des sujets (voir V-1.1.1.). Le point V-2.1.2 s'intéressera brièvement aux problèmes lexicaux non cibles, et nous tenterons, au point V-2.1.3, d'expliquer la différence entre les taux de résolution des problèmes lexicaux des deux parties du test. Finalement, la section V-2.1.4, quant à elle, tentera de comprendre pourquoi certains types de problèmes ont été résolus mieux que d'autres et observera le statut suffisant ~ insuffisant des séquences.

2.1.1 LE REPÉRAGE DES PROBLÈMES LEXICAUX CIBLES

Le tableau X (voir V-1.1.1) nous révèle le taux de repérage des différents problèmes lexicaux présents dans la première partie de notre test. À la lumière de ces informations, est-il possible d'identifier les types de problèmes lexicaux qui retiennent l'attention des étudiants universitaires? Nous ne le croyons pas.

Premièrement, il semble assez difficile d'établir des tendances entre les types de problèmes et le fait qu'il soit repérés ou non avec le nombre limité d'erreurs présentes dans la première partie de notre test. En effet, le texte soumis aux étudiants ne comportait que 17 problèmes lexicaux, qui respectaient les proportions révélées par les résultats de notre premier objectif (voir III-3.2). Par conséquent, certains types d'erreurs n'apparaissaient qu'une fois dans le texte. C'est le cas, par exemple, de la seule erreur d'orthographe lexicale du test (**d'avantage*), qui a été repérée par 75 % des sujets. Il serait plutôt hasardeux d'affirmer, sur la seule foi de cette observation,

que les erreurs d'orthographe figurent parmi les mieux repérées. Notre échantillon d'erreurs limité nous place dans cette situation pour au moins 7 types d'erreurs.

Par ailleurs, pour un même type d'erreur, nous observons de grandes différences dans les taux de repérage. Un exemple illustre bien ce problème : les deux erreurs de prépositions régies ont été relevées respectivement par 100 % et 12,5 % des équipes. Les quatre erreurs de sens présentent aussi des taux de repérage très variables : de 75 % à 12,5 %. Comment expliquer, par exemple, que 62,5 % des équipes ont relevé l'impropriété *corroborent* (pour signifier 'collaborent'), alors qu'une seule a mentionné l'utilisation de *conséquence* pour exprimer 'punition'? Est-ce le fait que le verbe CORROBORER est peu courant et que de nombreux sujets en ont vérifié le sens dans le dictionnaire « juste pour voir » ? Est-ce plutôt le fait qu'il existe une certaine parenté sémantique entre CONSÉQUENCE et PUNITION – une punition étant généralement la conséquence d'une faute – qui a en quelque sorte atténué le caractère erroné de cette impropriété? Ou encore que beaucoup d'enseignants utilisent CONSÉQUENCE au sens de 'punition' (après tout, peut-être se sentent-ils moins coupables « d'infliger des conséquences » que des punitions!) et que les sujets ont déjà été exposés à cette autre acception? Impossible de répondre... Nous ne croyons donc pas pouvoir indiquer si des problèmes d'un type particulier sont généralement relevés par les étudiants. Par exemple, nous ne pensons pas que le fait que les trois problèmes les plus repérés découlent de la transgression de propriétés de combinatoire grammaticale (préposition régie, caractère figé d'une locution, genre nominal) soit significatif.

Selon nous, c'est le problème lui-même, plutôt que le type de problème, qui aura une influence sur son repérage. Parfois, il se peut même que des facteurs extérieurs au problème ciblé fassent qu'il soit relevé par moins de sujets. Voici six critères qui, selon nous, ont joué sur le taux de repérage de différents problèmes présents dans le premier volet de notre test.

1. Le fait que le sujet soit familier avec l'unité lexicale concernée :

L'erreur d'orthographe **d'avantage* présente dans notre corpus a probablement été repérée par beaucoup de sujets (75 %) parce que DAVANTAGE est une unité lexicale fréquente et que la plupart des étudiants sont familiers avec son orthographe. Si nous avions plutôt choisi d'introduire **occurrence* dans notre corpus, il y a fort à parier que le taux de repérage aurait été beaucoup moins élevé. Un fait similaire a pu être observé avec le verbe CORROBORER utilisé pour signifier 'collaborer'. En effet, plusieurs équipes

se sont attardées sur ce verbe pour la simple raison qu'ils n'en connaissaient pas le sens; c'est donc un peu par accident que les sujets se sont aperçus que ce verbe ne correspondait pas au contexte dans lequel il était utilisé. Ainsi, nous pouvons supposer que des étudiants repéreront plus facilement des problèmes concernant des unités lexicales qu'ils connaissent et dont ils possèdent une certaine maîtrise. Par contre, une unité lexicale qui leur est inconnue pourra aussi susciter chez eux des questionnements qui les mèneront à l'identification de certains problèmes.

2. Le fait que le segment concerné produise un effet étrange parce qu'il ne « sonne pas bien » :

Pour les deux problèmes relevés par l'ensemble des équipes – *sont portés *de manquer et juste et 'bon milieu* –, l'identification s'est faite spontanément à la lecture des passages en question; tous les sujets ont eu la même réaction de surprise en les lisant et la plupart ont même buté sur le passage concerné tant ils ont trouvé que « Ça s'dit pas ben. [équipe 6, séq. 24] ». Il semble donc qu'un des facteurs dominants dans l'identification de problèmes lexicaux soit purement intuitif; lorsqu'il y a spontanément un jugement défavorable du locuteur natif, c'est qu'il y a forcément un problème. Ceci permettrait d'expliquer la grande différence entre les taux de repérage des deux problèmes de prépositions régies (100 % et 12,5 %) ; intuitivement, *être porté *de* est senti comme beaucoup plus problématique que *avoir de l'intérêt *envers*. Cette intuition provient peut-être du fait que la préposition À qui accompagne [être] PORTÉ est une préposition régie « pure » complètement vide de sens ; *être porté à* est donc perçu comme un bloc par les locuteurs. Le cas d'INTÉRÊT semble moins clair. Ce nom est plus sujet à accepter d'autres prépositions; nous n'avons qu'à penser à l'anglicisme **avoir de l'intérêt dans quelque chose* [**Il a peu d'intérêt dans les sports d'équipe.*] pour constater que le choix de la préposition qui accompagne INTÉRÊT est plus flexible. C'est sûrement pour cette raison que ce second problème de préposition n'a été relevé que par une seule équipe.

3. Le fait qu'un sujet soit plus sensible à un type de problème particulier :

Les deux cas d'anglicismes présents dans le test n'ont été repérés que par 25 % des équipes. Les deux anglicismes retenus – METTRE L'EMPHASE (sur) et OPPORTUNITÉ (= 'chance') – sont très fréquents en français québécois et il est facile de comprendre

qu'ils n'aient pas été jugés problématiques par la majorité des équipes. En fait, des 3 sujets ayant mentionné ces problèmes, un provient d'un milieu bilingue et a l'habitude de se méfier des anglicismes (« opportunity ». En anglais, c'est très utilisé. [équipe 4, séq. 12]), un connaissait spécifiquement ces anglicismes pour les avoir étudiés (moi, au cégep, j'avais appris que... [équipe 5, séq. 12]), et le troisième s'y est arrêté un peu par hasard, puisqu'il s'interrogeait sur à peu près tous les éléments du texte. Nous croyons donc qu'un anglicisme ne sera repéré que par des personnes sensibles à ce type d'erreurs parce qu'elles sont familières avec l'anglais, ou alors par d'autres qui auront appris spécifiquement les cas d'anglicismes les plus communs.

4. Le fait d'avoir déjà rencontré le problème ciblé ou un problème s'y apparentant :

Dans le cas de **différents idoles*, la plupart des sujets semblent s'être méfiés parce qu'ils avaient déjà été confrontés au même problème ou à un problème similaire (*C'est comme un autobus, on l'a jamais*. [équipe 1, séq. 17]). Forcément, lorsqu'un étudiant rencontre un problème auquel il se souvient s'être heurté, il sera plus en mesure d'identifier le « potentiel d'erreur ».

5. Le fait que le problème ciblé constitue véritablement une erreur ou soit plutôt considéré comme une maladresse lexicale :

Le seul cas qui n'a été repéré par aucune équipe est le cas de répétition : *en s'intéressant d'avantage* (sic) *aux intérêts*. Nous voyons quelques raisons à cela. La première raison est que ce problème ne constitue pas une erreur véritable, mais plutôt une maladresse. Par conséquent, certains sujets n'ont peut-être pas cru bon relever ce problème particulier, puisque le texte qui leur était soumis contenait plusieurs imperfections et que la consigne était de repérer les erreurs.

6. Le fait que l'attention du sujet soit détournée par d'autres éléments du contexte dans lequel apparaît le problème ciblé :

Une deuxième raison peut expliquer pourquoi le problème de répétition n'a été relevé par aucune équipe. En effet, une erreur d'une autre nature – la mauvaise orthographe de **d'avantage* – a en quelque sorte servi d'écran entre les deux éléments de la répétition et a sans doute détourné l'attention des sujets vers un problème plus

important. Il se peut donc que la simple position d'une erreur lexicale dans un texte ait une influence sur son repérage.

En résumé, nous croyons qu'il existe des facteurs, au-delà du type d'erreur, qui sont plus déterminants dans le repérage d'un problème lexical. Les plus importants semblent être l'intuition du locuteur et le fait de maîtriser ou non l'unité lexicale en jeu. Cette constatation nous rappelle la grande part d'arbitraire présente dans l'évaluation de la maîtrise lexicale, puisque les sujets sont évalués à partir d'un nombre limité d'unités lexicales qu'ils connaissent fort probablement à des degrés bien différents simplement parce qu'ils les ont rencontrées (ou non) au hasard de leurs lectures ou de leur formation scolaire. Il est impossible d'affirmer, sur la base d'une maîtrise satisfaisante d'un petit nombre de lexies choisies arbitrairement, qu'une personne possède une bonne maîtrise du lexique. Ceci renforce l'idée de Lépine (1995), qui affirmait qu'il est très difficile d'évaluer la maîtrise du vocabulaire à partir de courts textes. Il serait par contre intéressant de se pencher de nouveau sur la question du repérage des problèmes lexicaux avec un nombre plus élevé d'erreurs de chaque type pour constater si cette dimension a une influence sur le repérage.

2.1.2 LES PROBLÈMES LEXICAUX NON CIBLES

Nous avons mentionné au point V-1.1.1 que, sur les 136 séquences portant sur des problèmes lexicaux, 56 ne concernaient pas des problèmes cibles. Bien que nous désirions décrire les stratégies de résolution de problèmes lexicaux et que nous nous concentrons par conséquent sur les problèmes lexicaux réels (ou cibles), il est intéressant de jeter un coup d'œil rapide à ces 56 cas.

En premier lieu, il est important de mentionner que le fait de discuter de problèmes non cibles n'est pas mauvais en soi. En effet, nous avons défini un problème comme un questionnement relatif à toute dimension lexicale (voir III-3.3.1). Par conséquent, un élément correct d'un texte constitue un problème lexical dès l'instant où l'on s'y intéresse d'un point de vue lexical. Il n'est donc pas mauvais de s'interroger sur le sens d'une unité lexicale, par exemple, pour en vérifier l'adéquation avec le contexte dans lequel elle apparaît. Un correcteur expert doit d'ailleurs s'intéresser à plusieurs éléments corrects d'un texte pour parvenir à déceler ceux qui sont problématiques; cela fait partie du processus normal de correction. Par contre, une séquence portant sur un

problème non cible devrait être caractérisée par une absence de changement, puisque la consigne donnée aux sujets était de corriger les erreurs relevées. Il n'était donc pas pertinent, en regard de la consigne reçue, de changer un élément correct pour un autre élément correct. Bien entendu, le fait de changer un élément correct par quelque chose d'incorrect était encore pire, puisqu'il revient à ajouter des erreurs dans le texte à corriger! En analysant les 56 séquences portant sur des problèmes non cibles, nous serons plus à même de juger de leur pertinence.

Tous les problèmes non cibles discutés, sauf un, ne constituaient pas des erreurs au départ. La seule séquence non cible portant sur un problème réel est la séquence 16 de l'équipe 2, dans laquelle les sujets s'interrogeaient sur la cooccurrence du verbe IMPOSER avec le nom CONSÉQUENCE dans la phrase [...] *les professeurs doivent parfois imposer des conséquences aux élèves*. Plutôt que remettre en doute l'emploi de CONSÉQUENCE, les sujets tentaient de trouver un meilleur collocatif pour ce nom (« infliger des conséquences » ? [équipe 2, séq. 16]). Les 55 autres séquences non cibles concernaient donc des éléments corrects du texte. Pourtant, dans la majorité des cas (31 sur 56, soit 55,4 %), les sujets ont tout de même choisi d'effectuer un changement. La majorité des changements apportés ont laissé intact le statut du problème, c'est-à-dire qu'un élément correct était remplacé par un autre élément correct; il n'y a que 6 cas, sur les 31 changements effectués, qui ont transformé en erreur un élément correct au départ.

Une bonne partie des séquences portant sur des problèmes non cibles (44,6 %) consistaient à changer un terme pour un autre sans vraiment justifier la raison de ce changement (Moi, j'aime pas l'mot « proche », là. J'aurais mis « près de leur enseignant ». [équipe 6, séq. 21]), ou alors par des arguments esthétiques très subjectifs (J'trouve vraiment ça s'dit mal. [équipe 1, séq. 21]) ou des raisons sémantiques un peu nébuleuses ([remplace « c'est donc au professeur de... » par « il revient donc au professeur de... »] Parce que « c'est donc », ça se dit pas dans le sens que... il veut dire, parce que, y veut donner la responsabilité à quelqu'un et... le verbe « être » pour responsabiliser quelqu'un, c'est trop trivial. [équipe 4, séq. 24]). La plupart des autres séquences lexicales s'intéressant à des problèmes non cibles concernent des questionnements relatifs au sens d'unités lexicales présentes dans le texte (notamment LAXISME : [équipe 1, séq. 23], [équipe 3, séq. 10], [équipe 6, séq. 22]) ou à de faux pléonasmes (*étudier leur leçons*. Mais qu'est-ce que t'étudies, à part des leçons? C'comme dire « monter en haut ». [équipe 7, séq. 6]).

2.1.3 LA RÉOLUTION DES PROBLÈMES LEXICAUX DANS LES DEUX PARTIES DU TEST

Nous avons considéré, à la section V-1.1.2, tous les problèmes cibles traités pour établir le taux de résolution des problèmes à 80,5 %. Par contre, si nous analysons séparément les problèmes de la première partie du test (80 problèmes repérés) et ceux de la deuxième partie (120 problèmes ciblés), nous constatons un écart de plus de 3 % dans la résolution des problèmes. En effet, les problèmes correctement résolus représentent 82,5 % de ceux relevés dans la première partie³⁸, alors que 79,2 % ont été bien résolus dans la deuxième partie. À notre avis, cette différence somme toute peu importante s'explique simplement. En effet, les cas traités par les sujets dans la première partie l'ont été parce qu'ils constituaient véritablement des problèmes à leurs yeux. Dans la deuxième partie, par contre, les problèmes étaient en quelque sorte « imposés » aux sujets. Ainsi, les sujets se sont retrouvés confrontés à des problèmes qu'ils n'auraient de prime abord pas considérés comme tels; ils ont donc dû proposer des solutions à des problèmes qui, selon eux, n'auraient pas nécessairement demandé un changement. Cette situation les a donc poussés à apporter des modifications sans avoir bien cerné le problème en jeu, ce qui a occasionné de mauvaises résolutions. Nous pouvons donc supposer que les résultats de la première partie du test sont plus éloquents à l'égard de la résolution de problèmes : lorsqu'un problème lexical est repéré par des étudiants universitaires, il est bien résolu dans près de 83 % des cas. Étrangement, la proportion de séquences présentant des stratégies suffisantes est plus élevée dans la deuxième partie du test (71,7 %) que dans la première (57,5 %). En effet, comme les sujets se retrouvaient devant des cas qu'ils n'auraient pas nécessairement jugés problématiques dans la deuxième partie du test, nous aurions pu supposer que leurs stratégies de résolution pour ces problèmes auraient été moins souvent suffisantes que celles déployées pour des problèmes qu'ils auraient eux-mêmes repérés dans la première partie. Pourtant, les chiffres prouvent le contraire. Nous voyons deux raisons potentielles à cet écart assez important. La première est que les types de problèmes présents dans les deux parties n'étaient pas les mêmes et qu'il se peut simplement que les étudiants aient été plus habiles à expliquer certains types d'erreurs (nous pensons notamment aux problèmes de périphrase et de confusion entre

³⁸ Il est important de rappeler que seulement 58,8 % des problèmes présents dans cette partie du test ont été repérés. Par conséquent, le taux de résolution réel des problèmes pour la première section du test est de 48,5 %, ce qui signifie que plus de la moitié des problèmes demeuraient en fin de tâche.

parties du discours qui ont été expliqués correctement par la grande majorité des équipes et qui figuraient dans la deuxième partie du test). La deuxième hypothèse s'appuie sur le fait que les cas repérés dans la première partie sont ceux qui constituaient véritablement des problèmes aux yeux des sujets. Il se peut donc que les sujets aient mis moins d'efforts à analyser ces problèmes et qu'ils aient proposé spontanément des solutions. Ceci expliquerait le tût moins élevé de séquences suffisantes dans la première partie, puisque nous n'avons pas considéré suffisante une séquence présentant des solutions au problème sans analyse ou justification.

2.1.4 LA RÉOLUTION DES PROBLÈMES LEXICAUX PAR TYPES DE PROBLÈMES

Nous avons vu, à la section V-1.1.2, les résultats concernant la résolution des problèmes lexicaux cibles et le statut suffisant ou non suffisant des séquences qui ont permis de résoudre ces problèmes. La section suivante s'intéressera spécifiquement aux points du tableau XI (V-1.1.2) qui ont retenu notre attention.

Les deux problèmes relatifs à des propriétés inhérentes des unités lexicales présents dans le test (mauvais genre nominal et non-respect de la partie du discours) ont été correctement résolus par l'ensemble des équipes qui les ont traités. De plus, 93,3 % des séquences y étant associées ont été considérées suffisantes. En quoi consistaient ces stratégies et comment se fait-il que le problème ait été bien résolu dans tous les cas? Pour le problème lié au genre nominal, **différents idoles*, la stratégie de 6 équipes sur 7 fut d'aller vérifier le genre du nom IDOLE dans le dictionnaire; un des membres de la septième équipe connaissait le genre du nom IDOLE et s'est contenté de l'affirmer. Chercher le genre d'un nom dans le dictionnaire constitue une des utilisations normales que les gens en font et il est peu surprenant que les étudiants aient effectivement trouvé la réponse à leur questionnement dans les ouvrages. Quant à l'utilisation de ENTHOUSIASTE au lieu d'ENTHOUSIASME, cela semble correspondre à un problème connu des étudiants (J'fais tout le temps c't'erreur là ! [équipe 2, seq. 43] ; Moi, ça m'fait tellement rire le monde qui mélange ça, là. [équipe 6, seq. 40]) ; tous l'ont corrigé d'emblée et toutes les équipes, sauf une, ont fait référence à la partie du discours. Les deux problèmes de cette classe d'erreur retenus pour notre test étaient donc relativement simples et il est peu surprenant que les étudiants aient su non seulement les corriger, mais aussi justifier leurs décisions.

Les deux cas d'emploi de mauvaises prépositions ont aussi été correctement résolus chaque fois qu'ils ont été traités. Par contre, le taux de stratégies suffisantes ne s'élève qu'à 44,4 %. Dans le cas des prépositions régies, les changements ont parfois été effectués sans qu'aucune raison ne soit évoquée ([équipe 1, séq. 7] ; [équipe 8, séq. 11]), ou alors des raisons purement intuitives (Ben, j'sais pas, ça s'dit mieux, là. J'connais pas la règle, là, mais me semble... [équipe 3, séq. 3] ; Pasque... ça s'dit pas ! [équipe 5, séq. 5]). Seulement deux équipes ont mentionné le terme PRÉPOSITION (Au son, j'crois que « à », c'est une préposition qui va mieux. [équipe 6, séq. 10]) et les étudiants semblaient un peu dépourvus lorsqu'ils tentaient de trouver une réponse dans le dictionnaire (J'sais pas comment chercher ça là... C'est un peu problématique. [équipe 5, séq. 5]). En lisant les séquences concernant ce type de problème, nous constatons que les étudiants ne connaissent pas le phénomène de prépositions régies et ne disposent d'aucun élément pour justifier le choix d'une préposition particulière. Malgré tout, certains ont réussi à trouver dans le dictionnaire des exemples présentant l'emploi de la préposition régie, lorsqu'ils n'ont pas cherché réponse à leur problème sous l'article de la mauvaise préposition ([regarde l'article DE] Ça s'dit pas devant un verbe, là. En tout cas, ici « de, article partitif, s'emploie devant les noms ». Puis, préposition c'est la même chose là... un nom aussi là. [équipe 5, séq. 5]) ! Pourtant, ces informations sont assez simples à trouver dans les ouvrages lorsque l'on sait où chercher.

Les problèmes de sens ont attiré notre attention lors de l'analyse du tableau XI (V-1.1.2), puisqu'il y a correspondance presque parfaite entre le nombre de séquences présentant des stratégies suffisantes (39) et le nombre de problèmes bien résolus (40); nous avons donc voulu vérifier s'il les séquences suffisantes correspondaient à des problèmes bien résolus dans tous les cas. En regardant de plus près les séquences, nous sommes aperçus que 3 présentent des stratégies suffisantes et conduisent pourtant à de mauvaises résolutions. Des 3 séquences qui correspondent à ce cas de figure, deux portent sur le cas de création d'acception d'un vocable : *Il est étrange que des gens qui ont tous transité par le secondaire...* Dans les deux cas, les sujets mentionnent, après avoir cherché la définition de TRANSITER dans le dictionnaire, que le choix de ce terme n'est pas idéal, mais décident de le laisser ainsi. Par contre, il est intéressant de voir que ces deux équipes ont proposé le terme PASSER – qui convient mieux à notre avis – et l'ont écarté pour des motifs différents ; l'équipe 3 [séq. 21] a choisi de laisser TRANSITER faute de pouvoir prouver que PASSER était plus approprié, alors que l'équipe 1 [séq. 38] s'est ravisée après avoir écrit PASSER, parce qu'elle considérait que c'était trop « québécois » et qu'elle n'avait pas d'autre solution à proposer. Comme nous avons

nous-même classé ce problème lexical parmi les maladroites, il est tout à fait compréhensible que certaines équipes aient choisi de ne pas le corriger. Par contre, la troisième séquence au contenu suffisant débouchant sur une erreur est plus intéressante. Elle concerne l'emploi de CORROBORER dans la phrase *Faire en sorte que ces deux groupes corroborent est un réel défi [...]*. Après que sa collègue lui a proposé de remplacer *corroborent* par *collaborent*, l'autre membre de l'équipe 6 admet qu'il ignore le sens de CORROBORER, en cherche la définition dans le dictionnaire, la lit à voix haute avant de déclarer « Ouais. C'est correct. [équipe 6, séq. 13] ». Voici un exemple parfait de séquence qui, bien que l'ensemble des éléments nécessaires à la résolution du problème soient présents, débouche sur une mauvaise solution.

Les problèmes de collocations ont été résolus dans 63,2 % dans cas où ils ont été traités. Il est intéressant de regarder les séquences concernant ces problèmes, puisque nous savons que ce phénomène linguistique n'est pas souvent enseigné. En analysant ces séquences, nous constatons que, malgré tout, les sujets semblent avoir une connaissance intuitive du phénomène. Ce commentaire d'un membre de l'équipe 7 témoigne de cette intuition : « Tu peux pas « donner un service », c'est tu « rends un service ». Tu comprends? Faque on peut peut-être pas « donner l'accès », on peut peut-être « offrir l'accès ». Je l'sais pas... [équipe 7, séq. 21] » Lors de l'analyse des problèmes, par contre, nous sentons que les sujets manquent de ressources pour bien expliquer ce à quoi ils sont confrontés. C'est surtout lorsqu'ils doivent trouver une solution après avoir constaté la présence d'un mauvais collocatif qu'ils se sentent dépourvus. Plusieurs tentent d'expliquer le problème en allant voir la définition du collocatif dans le dictionnaire (Est-ce qu'on peut employer une autorité? J'sais pas, han? On regarde-tu la définition d' « employer » ? [équipe 6, séq. 12]), ce qui les conduit souvent à accepter le choix d'un terme qu'ils avaient spontanément rejeté (« Employer, faire usage de ». Bon, ok. Y'ont raison, là. [équipe 5, séq. 7]). Par contre, certains ont exprimé leur difficulté à utiliser le dictionnaire pour les aider dans des cas semblables (Dans l'fond, sur quoi qu'on pourrait vérifier? Ça serait sur « autorité »... Mais y voudront pas... y diront pas... euh. J'vas vérifier pareil, là, mais j'pense pas qu'y disent... [équipe 2, séq. 12] ; J'essaie d'voir où est-ce qu'on pourrait trouver ça. [prend le dictionnaire] Ben, j'pense pas qui nous l'indiquent, mais... [équipe 2, séq. 40] ; Ben, c'pasque... y nous disent pas qu'est-ce... tsé, c'est quoi le... le verbe qui peut aller avec « accès », là. [équipe 2, séq. 30]). Une meilleure connaissance du phénomène de collocation pourrait sans aucun doute aider les étudiants à mieux aborder ce type de problème, notamment en se fiant davantage à leur intuition de

locuteur et en cherchant les collocatifs les plus communs dans l'article de dictionnaire de la base de la collocation.

Le cas de non-respect du mode verbal imposé par l'unité lexicale BIEN QUE a aussi retenu notre attention. Même s'il a été résolu correctement par trois équipes, aucune des cinq séquences s'y intéressant ne présente des stratégies suffisantes. Les cinq équipes ayant traité ce problème ont suggéré de remplacer le verbe au présent par un subjonctif, dont 2 de façon explicite (Ici, tu connais le subjonctif? On devrait pas mettre « bien que le fouet ne soit plus utilisé ». [équipe 6, séq. 6] ; Pis d'ailleurs « Bien que le fouet ne soit plus », pasque c't'avec le sub... jonctif, c'tu ça ? [équipe 2, séq. 6]). Deux des trois équipes qui ont effectué la correction l'ont fait de façon intuitive, sans pouvoir justifier (« Ne soit », pas « n'est » [équipe 1, séq. 4] ; Je l'sais pas pourquoi qu'y faut l'dire de même, mais c'est plus beau. [...] Tsé, c'est comme les trucs pour les homophones pis toute ça... Je l'sais pas, ça vient tu seul. [équipe 7, séq. 3]). Certaines équipes ont justifié le changement de mode par la présence du *que* dans la phrase (C'est l' « que » qui m'dit qu'on devrait mettre un subjonctif. [équipe 6, séq. 6] ; Ouais, c't'à cause du « que », mais euh... [équipe 2, séq. 6]). Par contre, un membre de l'équipe 2 semblait percevoir que la présence de la locution, et non seulement du *que*, expliquait l'emploi du subjonctif (Le pourquoi c'est un peu loin, là, mais « bien que »... ça implique une... [équipe 2, séq. 6]). Les deux équipes qui ont décidé de laisser le mode indicatif ont pris leur décision faute d'arguments solides pour appuyer un changement (Humm... J'sais pas si ça s'rait pertinent d'mettre le subjonctif. On peut l'laisser au présent. [équipe 6, séq. 6] ; C'est juste... Ben, les deux temps d'verbes, c'est juste c'est euh... c'est un niveau d'vocabulaire plus haut « ne soit » que « n'est plus », mais c'est la même chose. [équipe 4, séq. 4]). Il semble donc que les étudiants aient perçu intuitivement le problème de mode verbal dans ce cas-ci, mais aucun n'a réussi à l'expliquer. Leur manque de ressource face à un tel problème découle sûrement du fait qu'ils n'ont pas su identifier la locution BIEN QUE et qu'ils n'ont pas cherché dans le dictionnaire; en effet, les deux ouvrages de référence dont ils disposaient mentionnaient, sous l'entrée de BIEN, la locution BIEN QUE et spécifiaient qu'elle doit être suivie du subjonctif. Aucune équipe n'a eu recours au dictionnaire pour ce problème.

Les deux problèmes liés au non-respect de la structure actancielle ont été traités de façons très différentes par les équipes. Le premier problème n'a été résolu que par une seule équipe, alors que 7 équipes sur 8 ont correctement solutionné le second. Pourtant, seules 12,5 % des séquences relatives à ces problèmes présentaient des stratégies suffisantes. Le premier problème, – *Un jour, j'ai décidé de donner l'accès à mon frère.* – n'a été bien résolu que par une seule équipe (« d'en donner l'accès ». Pasque

sinon c'est... ça dit pas donner l'accès de quoi. [équipe 6, séq. 27]). Une seule autre équipe a souligné l'absence du complément du nom ACCÈS, mais a proposé une mauvaise correction (Si t'as pas de pronom qui rappelle à quoi tu donnes l'accès, ta phrase a pas de sens. [...] On pourrait dire « d'y en donner l'accès », ou juste « d'y donner l'accès ». Les deux sont bons, les deux sont très bons. [équipe 4, séq. 27]). La plupart des équipes se sont plutôt interrogées sur le choix du collocatif DONNER avec le nom ACCÈS (Bien, tu peux pas donner l'accès à quelqu'un. Tu peux donner une clé d'accès, là, mais tu peux pas donner l'accès. Me semble. Ça s'donne pas un accès. [équipe 7, séq. 21] ; On donne pas un accès, on ouvre un accès, non? [équipe 8, séq. 36]) ou encore sur la présence du « L » apostrophe dans l'expression *donner l'accès* (« donner accès » au lieu de « donner l'accès ». Me semble que « donner-r'accès », ça sonne bien. [équipe 5, séq. 26] ; Ouais, mais genre, t'enlèves le « L » apostrophe. Tu peux pas « donner l'accès ». [équipe 1, séq. 34]). Le fait que les sujets s'attardent sur des éléments périphériques nous laisse supposer qu'ils n'auraient sans doute pas considéré le problème soumis comme réellement problématique. Par conséquent, les stratégies ici déployées ne sont probablement pas très révélatrices des stratégies véritablement utilisées dans des cas de non-respect de la structure actancielle d'une unité lexicale.

Le second problème de cette nature (*elle lui informe ses déplacements*), par contre, est plus éloquent en ce qui concerne les stratégies, ou plutôt l'absence de stratégies, des étudiants. En effet, ce problème a été correctement résolu par l'ensemble des équipes, sauf une (« elle lui informe de ses déplacements » [...] Parce que c'est un... COD. Pis y faut un « de ». [équipe 4, séq. 35]). Les 7 autres équipes ont spontanément proposé la structure correcte, mais une seule a su expliquer sa décision de façon satisfaisante. Certains sujets ont tout de même mentionné que le verbe INFORME nécessite la préposition DE pour introduire un de ses compléments (Parce que « de » introduit une idée en faite. C'est « elle l'informe de quelque chose ». [équipe 8, séq. 44] ; Pis « t'informes de quelque chose ». [équipe 5, séq. 34]). C'est le remplacement du *lui* par *l'* que les étudiants avaient beaucoup plus de difficulté à expliquer. Un des membres de l'équipe 7 a tout de même diagnostiqué le problème de façon assez claire, mais sans réussir à l'analyser plus en détails (Ben, « lui informer »... Y'a comme une faute de pronominalisation. Ça s'peut-tu? [équipe 7, séq. 29]). Un autre sujet, dont la séquence a été jugée suffisante, détenait la solution, mais sans disposer du métalangage qui lui aurait permis de l'expliquer de façon élégante (« elle l'informe », c'est ça? Pasque qu'on informe quelqu'un, on n'informe pas « à » ou « de » quelqu'un, là. Tsé c'est comme... j'sais pas comment dire ça, mais en tout cas. [équipe 2, séq. 38]). Étrangement, 3 équipes ont justifié

le « L » apostrophe par des raisons morphophonologiques (Ben, « lui », c'est pas que le mot « informe » commence avec un « l », faque tu peux pas dire « lui ». [...] si l'mot « informe » commence avec une voyelle, y faut que tu mettes un « L » apotrophe. [équipe 8, séq. 44] ; Ah! Pasque y'a deux « l » de suite... C'tu ça? Y'a deux voyelles... [équipe 3, séq. 25]). Devant leur incapacité à pouvoir justifier convenablement leur correction, certaines équipes ont eu l'idée ingénieuse de proposer un synonyme approximatif d'INFORMER qui correspondait à la structure erronée (« elle lui décrit ses déplacements », ça serait bon. [équipe 6, séq. 35]). Dans le cas de ce deuxième problème de structure actancielle, donc, nous pouvons conclure que la très grande majorité des sujets avaient une connaissance de la structure correcte régie par le verbe INFORMER ; ce sont plutôt les règles de pronominalisation du premier complément qui leur ont posé problème, ce qui relève davantage de la syntaxe du français que d'une connaissance purement lexicale.

Le problème d'utilisation du mauvais auxiliaire dans *il avait même monté dans le mauvais autobus* a été résolu correctement par 4 équipes. Les 4 autres équipes ont soit remplacé le verbe MONTER par EMBARQUER (Ben, c'pas « monter », là. « embarquer ». Pasque « monter », c'est pas l'même sens. [équipe 3, séq. 20]), soit changé le plus-que-parfait pour le passé composé (« il a même monté dans le mauvais autobus ». [...] C'est déjà le passé composé. Pas obligé de mettre « avait ». [équipe 8, séq. 39]). Par contre, une des équipes ayant changé MONTER par EMBARQUER avait aussi ciblé le problème de l'auxiliaire (J'sais pas pourquoi, mais j'mettrais « avait embarqué » pis « était monté ». [équipe 1, séq. 37]). Les équipes ayant bien résolu le problème ont généralement fourni des explications suffisantes (On dit « je suis monté », pas « j'ai monté ». [équipe 6, séq. 30] ; Parce que c'est un verbe que au passé composé, il se conjugue avec « être », pas avec « avoir ». [équipe 4, séq. 30] ; Le verbe, là, quand que c'est conjugué, là, comment que t'appelles ça? L'auxiliaire!... employé est donc « être ». [équipe 7, séq. 24]). Par contre, nous avons choisi le verbe MONTER spécialement parce que c'est un verbe qui, selon son caractère transitif ou intransitif, pourra être conjugué avec l'auxiliaire AVOIR ou l'auxiliaire ÊTRE. Une seule équipe s'est heurtée à ce problème lorsqu'elle a rencontré un exemple d'emploi avec AVOIR en cherchant dans le dictionnaire (« monter : parcourir de bas en haut, elle a monté l'escalier... » Hey! Ça tient pus... [équipe 7, séq. 24]). Après avoir longtemps hésité, et sans avoir compris que MONTER pouvait être utilisé avec les deux auxiliaires, les membres de l'équipe 7 ont tout de même choisi de se fier à leur intuition (Moi, j'étais sûre que c'était « était », mais là, y m'disent que j'sais pas ça. J'suis

frustrée. [...] Ah... Ça s'dit mal « il avait même monté dans le mauvais autobus ». [remplace « avait » par « était »] Ostie de dictionnaire poche! [équipe 7, séq. 24]).

De façon globale, nous pouvons retenir de ce survol des séquences produites par les étudiants :

1. que de nombreux problèmes ne sont pas repérés (41,2 %);
2. que les problèmes traités sont généralement bien résolus;
3. que les étudiants ignorent la richesse des informations contenues dans les dictionnaires;
4. qu'ils ont une connaissance intuitive de certains phénomènes lexicaux sans arriver à bien les définir (prépositions régies, collocation, etc.) ;
5. qu'ils se fient peu à leur intuition de locuteur, ce qui les pousse à opter pour des solutions qu'ils auraient spontanément écartées.

2.2 LES COMMENTAIRES

Les sections qui suivent s'intéresseront à l'analyse des commentaires constituant les séquences. Par l'analyse du contenu de ces commentaires, nous serons plus à même de comprendre les stratégies mises en œuvre par les étudiants dans la résolution de problèmes lexicaux. La section V-2.2.1 commentera certains chiffres présentés à la section précédente. La section V-2.2.2 s'intéressera aux types de commentaires les plus intéressants. Finalement, la section V-2.2.3 s'interrogera sur la pertinence des stratégies utilisées pour chacun des types de problèmes de notre test.

2.2.1 LES RÉSULTATS GLOBAUX

La spontanéité des commentaires :

Le fait que 88,5 % des commentaires aient été émis spontanément par les sujets nous laisse croire que notre analyse rend vraiment compte des stratégies effectivement déployées par les étudiants confrontés à des problèmes lexicaux. En effet, nous aurions pu douter du caractère représentatif des stratégies observées si elles avaient été davantage orientées par l'intervieweur en cours d'entrevues. Seulement 176 commentaires ont été suscités par des interventions de l'intervieweur. Parmi ceux-ci, la très grande majorité (164) ne visaient qu'un approfondissement d'un point déjà soulevé par les sujets; ces 164 interventions de l'intervieweur ne contenaient aucune piste d'analyse ou de solution et n'avaient pour but que d'avoir accès aux

raisonnements effectués de façon implicite par les sujets. Il n'y a que 12 commentaires – moins de 1 % du total – qui ont été émis après que l'intervieweur a proposé une piste d'analyse ou fourni un élément de réponse; ce sont les seuls qui ne témoignent probablement pas des stratégies qu'auraient naturellement mises en œuvre les étudiants. Bien que la situation même d'entretien métagraphique soit un peu artificielle et que toute réflexion à voix haute puisse paraître peu spontanée, nous pouvons considérer que plus de 99 % des verbalisations codées rendent compte des réflexions réelles des étudiants devant un problème lexical, puisqu'elles n'ont pas été influencées par l'intervieweur.

Le statut approprié ou non approprié des commentaires :

Nous avons vu au point V-1.2.1 que 962 verbalisations (62,8 %) ont été jugées appropriées. Nous avons mis en parallèle ce chiffre avec le taux de résolution des problèmes présenté à la section V-1.1.2, qui est de 80,5 %. Comment se fait-il que 80,5 % des problèmes soient bien résolus alors que seulement 62,8 % des commentaires composant les séquences sont appropriés? D'après nous, cela s'explique par le fait qu'une séquence non suffisante peut tout de même conduire à une bonne résolution du problème. Ainsi, une séquence contenant plusieurs commentaires inappropriés pourra tout de même conduire à résoudre correctement le problème traité. D'ailleurs, en comparant le taux de commentaires appropriés (62,8 %) avec le taux de séquences (66,0 %) suffisantes de la section V-1.1.2, nous observons qu'ils sont très proches. Par ailleurs, l'écart entre le taux de problèmes bien résolus et le taux de commentaires appropriés peut être expliqué par une autre hypothèse. Il se peut en effet que les séquences menant à une bonne résolution du problème contiennent moins de commentaires et que ces derniers soient généralement appropriés; ainsi, un nombre restreint de commentaires appropriés pourrait conduire à résoudre correctement le problème. D'autre part, il se peut que les séquences correspondant à des problèmes mal résolus contiennent davantage de commentaires dont la plupart seraient inappropriés. Cette hypothèse rendrait compte du fait que les problèmes rencontrés ont été généralement bien résolus même si près de 40 % des commentaires émis par les sujets sont inappropriés.

Les étapes de la résolution du problème :

Il est encourageant de voir que l'étape d'analyse du problème est celle qui a entraîné le plus de commentaires (761, soit 49,7 %). En effet, cette information témoigne du fait que les étudiants ont tenté de comprendre et d'expliquer les problèmes auxquels ils étaient confrontés. Nous avons d'ailleurs vu à la section II-3.2 que la phase d'analyse du problème est cruciale dans sa résolution et que les experts consacrent généralement beaucoup d'attention à cette étape. De plus, ces commentaires des étudiants se sont révélés pertinents dans 61,5 % des cas, alors les verbalisations inappropriées ne représentent que 21,7 %³⁹ des cas. Ceci indique que les sujets ont proposé des analyses dont beaucoup d'éléments étaient justes en regard du problème traité.

Évidemment, beaucoup de commentaires (414, soit 27 %) ont aussi porté sur la proposition et l'analyse de solutions aux problèmes discutés. Si plus du quart de ces commentaires étaient inappropriés (110 cas, soit 26,6 %), ils ne correspondent pas exclusivement à des problèmes mal résolus. En effet, il arrivait souvent que les sujets proposent de mauvaises solutions avant de les écarter ([...ces 92 pages nous récitent] « nous récitent » ? Ça peut pas réciter, des pages. « Nous indiquent » ? Non, peut-être pas « indiquer » dans l'sens d'une histoire, là. [équipe 2, séq. 39]).

Le nombre de commentaires concernant le repérage des problèmes est peu élevé (117, soit 7,6 %). Cette situation s'explique facilement. En effet, comme les problèmes étaient ciblés dans la deuxième partie du test, il est normal que les commentaires de repérage aient été plus rares. Quatre-vingts et trois dixièmes pour cent (80,3 %) des commentaires de repérage concernent la première partie du test. Ceux appartenant à la deuxième partie ont été émis lorsque les sujets ne voyaient pas de problème avec l'erreur ciblée (Je sais pas... moi j'vois pas. Tu vois-tu ? [équipe 2, séq. 40]), ou alors lorsque que le segment en caractères gras était plus long et que les étudiants tentaient de cerner où se situait précisément le problème (Est-ce que ça se dit « monter dans l'autobus » ? [équipe 7, séq. 24]).

Il est intéressant de remarquer que si les quatre étapes de résolution du problème sont souvent présentes à l'intérieur d'une séquence, elles n'apparaissent pas toujours dans l'ordre auquel nous nous serions attendu. Les commentaires relatifs au repérage se trouvent souvent en début de séquence, mais les autres étapes apparaissent dans des ordres différents. Il est fréquent, notamment pour les problèmes les plus repérés et ceux étant les mieux corrigés, que l'étape de décision précède l'analyse. Il arrive aussi

³⁹ Le reste des commentaires relatifs à l'analyse (16,8 %) ont été codés *indéfini*.

souvent que les sujets propose une correction à un problème avant même de l'expliquer ou de l'analyser. Nous croyons que ceci s'explique par le fait qu'une bonne partie de réflexions sont faites de façon implicite; ainsi, lorsqu'il propose une solution, le sujet a probablement déjà analysé le problème dans sa tête. C'est la situation d'entretien métagraphique qui l'oblige à exprimer les étapes qui précèdent la prise de décision (voire à en prendre conscience) pour des problèmes qui lui semblent évidents. Dans le cas de problèmes moins clairs ou tranchés, les sujets ont dû analyser davantage pour être à même d'élaborer des solutions, ce qui explique que l'ordre des étapes corresponde mieux au modèle proposé dans notre cadre conceptuel (voir II-3.3.3).

L'emploi du métalangage :

Nous avons vu plus haut que les sujets ont peu eu recours au métalangage dans la résolution des problèmes lexicaux; seuls 12,4 % des commentaires contenaient des termes propres à l'étude de la langue. On remarque quand même 248 occurrences d'éléments de métalangage dans nos transcriptions, et il est encourageant de constater qu'ils ont été bien utilisés dans 94,4 % des cas.

Les termes les plus fréquents dans notre corpus sont des éléments assez communs du métalangage : 71 (28,6 %) sont des noms de parties du discours (*verbe, adjectif, préposition*, etc.), 32 (12,9 %) concernent le genre (*masculin, féminin*) ou le nombre (*pluriel, singulier*) et 21 (8,5 %) désignent des concepts relatifs aux temps et mode verbaux (*subjonctif, infinitif, présent, passé composé*, etc.). Certains termes du métalangage relatif à l'étude du lexique ont tout de même été utilisés à quelques reprises : *anglicisme* (13), *synonyme* (12), (*langage*) *familier* (7), *connotation* (6), *expression* (5), *redondant* (4), *définition* (3), *niveau de langue* (2), *langage populaire* (2), *vocabulaire* (2). Parmi les hapax (termes n'apparaissant qu'une seule fois dans le corpus), nous relevons certains termes propres au lexique : *québécoisme, niveau de vocabulaire, régionalisme, orthographe, archaïsme, pléonasme, paronyme, homophone, langage courant*.

Nous constatons donc que les étudiants possèdent (et maîtrisent) un certain métalangage. Par contre, certains termes semblaient parfois leur manquer : C't'une expression qui va ensemble. [équipe 2, séq. 31]; Non, y'a pas de... de... C'est « davantage » en un mot. [équipe 7, séq. 12]; Moi, j'aurais dit « portés à », mais est-ce que c'est vraiment une bonne euh... [équipe 2, séq. 8]; « que revient la tâche »... un chapeau, c't'une « tache », ou euh... c'est quel ? [équipe 1, séq. 28]; Ah, comment ça s'appelle, là, un truc formé d deux mots comme ça? [équipe 2, séq. 35].

La dernière remarque que nous désirons faire à propos du métalangage concerne la confusion fréquente entre *langage familier* et *français québécois*. En effet, plusieurs sujets semblent penser qu'un terme propre au français québécois et nécessairement familier et doit par conséquent être « corrigé ». Voici quelques extraits qui témoignent de cette confusion : ... un dictionnaire euh... de la langue québécoise. Un dictionnaire joual. [équipe 4, séq. 14]; Ça s'dit pas. « tant qu'à moi », c'est pas du pur québécois? [équipe 8, séq. 48]; c'comme du « parlé », mais pas du « écrit », là. C'comme un Québécoisme... québécoisme? [équipe 1, séq. 36]; ben ça s'dit en québécois. Encore là... Tu l'dis quand tu parles, mais... J'pas sûre quand tu l'écris. [équipe 1, séq. 39].

2.2.2 LES TYPES DE COMMENTAIRES

À la lecture du tableau XIV (V-1.2.2), nous avons été surpris par la grande variété dans les commentaires émis par les sujets. Comme il nous sera impossible de discuter tous les types de commentaires, nous ferons ressortir les points qui nous semblent les plus intéressants dans cette analyse des commentaires.

Les commentaires les plus fréquents sont ceux véhiculant une **proposition de solution au problème traité dans le cadre d'un raisonnement** (280, soit 18,3 %). Cette constatation n'est pas d'un grand intérêt si nous considérons que la tâche des sujets consistait à discuter des problèmes rencontrés et à proposer des solutions pour y remédier. Ces commentaires étaient donc prévisibles et ne font que répondre à la tâche que les sujets avaient à effectuer; comme nous l'avons mentionné dans notre méthodologie (voir III-3.4.2), ils font partie du processus normal de résolution de problèmes lexicaux. Par contre, il est intéressant de contraster ces commentaires avec ceux dénotant une correction catégorique exempte de tout raisonnement et de toute justification, qui représentent 3,9 % du total des commentaires. Nous constatons que ces corrections catégoriques correspondent dans 80 % des cas à une résolution correcte du problème traité. Évidemment, les séquences dans lesquelles apparaissent ces commentaires ont dans la plupart des cas été considérées comme insuffisantes. Nous pouvons donc supposer que les étudiants ont recours à une correction non appuyée dans les cas où ils ont une bonne connaissance du problème ou, à tout le moins, une forte intuition sur la façon de le résoudre ([... ceux-ci sont plutôt portés de manquer leurs cours...] Euh... « sont plus portés à manquer ». [remplace « de » par « à »] [équipe 4, séq. 6]). Il y a même lieu de croire que, pour ces problèmes particuliers, l'analyse est faite

très rapidement et de façon implicite, de sorte que le sujet n'en a peut-être même pas conscience.

La consultation d'ouvrages de référence figure aussi parmi les stratégies de résolution de problèmes les plus souvent mises en œuvre par les sujets (137 cas, soit 8,9 %). Cette stratégie semble d'ailleurs efficace, puisque les séquences dont elle fait partie correspondent à des problèmes bien résolus dans 75,2 % des cas. Près de la moitié des recours au dictionnaire (environ 46 %) avaient pour but de vérifier le sens de l'unité lexicale cherchée et ont conduit à une bonne résolution 85,7 % des fois, c'est-à-dire que l'unité lexicale a été changée si son sens ne convenait pas au contexte et que les sujets l'ont laissée s'ils la considéraient appropriée. Il semble donc que les étudiants savent assez bien se servir d'un dictionnaire quand il s'agit de problèmes de sens. Les 10 recours aux ouvrages de référence dans le but de vérifier des propriétés de combinatoire grammaticale correspondent tous à des problèmes bien résolus; il faut cependant souligner qu'il s'agissait dans la majorité des cas de vérifier le genre du nom IDOLE ([différents idoles] Idole, nom féminin. Ah, ah ! [ajoute un « E » à « différents »] [équipe 1, séq. 17]). L'équipe 2, qui s'interrogeait sur la possibilité de mettre LEÇON au pluriel a réussi à trouver la réponse dans la zone d'exemples de l'article (Les leçons, ça s'compte-tu ? Hmmm... J'pas sûre. [prend le *Petit Robert*] Ben « leçons et devoirs », eux autres, ici, y l'mettent au pluriel, mais j'comprends dans c'sens-là, là. [...] Oui ! « apprendre, étudier, repasser, revoir ses leçons ». Ok. [équipe 2, séq. 11]).

Les sujets ont eu recours 12 fois au dictionnaire pour résoudre des problèmes de registre de langue. Il est étonnant de constater qu'il n'ont mentionné qu'une seule fois que l'unité lexicale appartenait au registre familier, alors que le dictionnaire l'indiquait clairement; les sujets ont même, dans plusieurs cas, lu intégralement le passage de l'article correspondant à la lexie traitée en omettant exclusivement l'information relative au registre de langue ([... les élèves déprimeraient moins...] « Je l'ai trouvé très déprimé...être abattu, démoralisé ». Ouais... [équipe 2, séq. 26]). Si les étudiants semblent conscients de l'existence des registres de langue – ils y ont fait référence 17 fois avec des termes de métalangage –, ils semblent ignorer qu'ils peuvent trouver des informations relatives à cette dimension de l'unité lexicale dans les dictionnaires.

Dans le cas des problèmes de prépositions régies, la vérification dans le dictionnaire a été « payante » lorsque les sujets ont pris le temps de lire les exemples (Son intérêt pour le sport! Tsé, tsé! [équipe 1, séq. 26]; Etre porté à! Ouais, c'est ça. C'est c'qu'on a dit. Avoir l'habitude de. Il est porté. [équipe 2, séq. 8]), mais s'est révélée un cauchemar lorsqu'ils ont tenté de trouver la réponse en se référant à l'article de la préposition ([... sont plutôt

portés de manquer leurs cours...] Ça s'dit pas devant un verbe, là. En tout cas, ici « de, article partitif, s'emploie devant les noms ». Puis, préposition c'est la même chose là... un nom aussi là.). Pour les problèmes de collocations, l'emploi du dictionnaire s'est révélé peu efficace dans la mesure où les sujets ont souvent cherché à valider une collocation douteuse en se référant à l'article du collocatif (« [les enseignants] doivent employer leur autorité ». [...] Ben, quoique là, à la limite, j pense que... c'est peut-être pas, justement, le bon choix d'mot. [...] Est-ce que c't'une faute... [regarde dans le Multidictionnaire] « Employer, faire usage de ». Bon, ok. Y'ont raison, là. [équipe 5, séq. 7]).

Nous constatons donc que, globalement, les étudiants sont conscients que le dictionnaire est un outil précieux dans la résolution de problèmes lexicaux, mais qu'ils ne semblent pas connaître la richesse d'information qu'il contient, ou à tout le moins la façon d'y accéder efficacement. En connaissant davantage la structure des articles et l'information qui y est présentée, ils auraient été en mesure de trouver des réponses à plusieurs questions. Nous désirons aussi rappeler que, même si l'utilisation au dictionnaire figure parmi les stratégies les plus fréquentes et les plus efficaces, les étudiants n'y ont eu recours que 137 fois sur 321 séquences (42,7 % des cas); une utilisation plus systématique aurait peut-être permis de résoudre davantage de problèmes.

Le recours au sens d'une unité lexicale sans se référer au dictionnaire est aussi une stratégie très fréquente dans notre corpus (134 cas, soit 8,8 %). Évidemment, cette stratégie a surtout été utilisée pour les problèmes de sens (51,9 % des cas) et s'est avérée très efficace pour ce type de problèmes, puisqu'elle correspond à une bonne résolution dans 87 % des cas. Les « définitions » proposées par les étudiants consistaient souvent en un simple synonyme de l'unité lexicale traitée (Non, « transiger », c'est « négocier », ça a pas rapport. [équipe 4, séq. 31]; « cor... corroborent » ? Ouin, comme approuver, là. [équipe 1, séq. 12]), mais nous avons observé certaines tentatives de définitions plus élaborées (Ça veut dire quoi « laxisme »? Euh... ben c'est comme, y'a trop de... Y laisse comme trop d'espace aux jeunes, tsé... sont trop permissifs. [équipe 3, séq. 10]; C'est genre euh... tout accepter, j'crois. Tout prendre comme bon. Mais on peut voir. [équipe 6, séq. 22]). Dans certains cas, les sujets ont fait ressortir une composante de sens en particulier (Tu peux pas explorer un objet, genre, encore là, t'explores un lieu, là. [équipe 1, séq. 45]; Ben, « réciter », c'est de vive voix, ça. C'est ça, c'est à haute voix, faque ça peut pas marcher. [équipe 3, séq. 26]; Ok. « emmener » c't'une personne, « emporter » c't'une chose. [équipe 6, séq. 26]).

Selon notre analyse, les sujets ont eu **recours à leur jugement de locuteur** dans 71 cas (4,6 %). En fait, nous croyons que cette stratégie a été employée beaucoup plus

souvent, mais que notre grille ne nous a pas permis de le mettre en évidence. Parmi les *types de commentaires*, trois codes nous servaient à identifier les commentaires véhiculant un jugement : *recours au jugement du locuteur*, *jugement global* et *jugement esthétique* (voir tableau V, III-3.4.2). En analysant les 235 commentaires codés *jugement global*, nous constatons que plusieurs, voire la plupart, véhiculent aussi un jugement de locuteur (« Pour en terminer avec la discipline », c'est clair qu'ça s'dit pas, là. [équipe 1, séq. 20]; Ça fait journaliste. Mais ça s'dit très bien « question de ». [équipe 4, séq. 32]; Ben, on dirait que l'autre ça se dit pas. Ça, on dirait que j'suis sûr que ça s'dit. [équipe 5, séq. 24]; T'as raison, ça se dit pas. [équipe 7, séq. 18]). Les seuls jugements globaux qui ne constituent pas à proprement parler un jugement de locuteur sont ceux qui arrivent en fin de séquence et découlent d'un raisonnement (Ben, « récitent », avec la définition, c'est bon. [équipe 7, séq. 32]; Mais, ici, dans le Petit Robert, on dit euh... On dit que ça ben d'allure « passer ». [équipe 5, séq. 30]; [après avoir découvert que l'unité lexicale EMMENER existe] Ouais, ben on apprend quelque chose, mais c'toujours pas bon pareil. [équipe 3, séq. 16]). Par ailleurs, nous avons aussi codé 49 commentaires véhiculant un jugement esthétique (commentaires diagnostiques), dont seulement 9 correspondent à des verbalisations ayant aussi reçu le code *recours au jugement du locuteur*. Encore une fois, nous croyons que ces deux codes auraient toujours dû apparaître ensemble, puisque tout jugement esthétique constitue aussi un jugement de locuteur (Mais « à mon avis », c'est plus joli. [équipe 1, séq. 46]; C'est juste parce que bon... c'est plus beau. [équipe 7, séq. 3]; J'avoue que « déprimerait », ça sonne pas bien. [équipe 8, séq. 27]; [sourir] J'trouve ça laid, moi, « il avait monté »! [équipe 7, séq. 24]). Donc avons donc recompté le nombre de commentaires véhiculant au moins un type de jugement : le total est de 330. Nous pouvons donc dire que, de tous les commentaires codés, 21,6 % contiennent un jugement, que ce jugement soit d'ordre esthétique, qu'il découle d'une intuition de locuteur ou qu'il soit fait à la suite d'un raisonnement.

L'énonciation de « règles » relatives à une unité lexicale est un type de commentaires que nous avons ajouté à notre grille de codage après avoir constaté que les sujets ont fait appel à cette stratégie à quelques reprises (16, soit 1,4 % des cas). Il est intéressant de constater qu'une grande partie de ces commentaires concernent les paronymes EMMENER, EMPORTER, AMENER et APPORTER (J'pense c'est ça, « t'emmène quelqu'un » pis « t'amène quelque chose », quelque chose du genre... ou l'inverse!. [équipe 1, séq. 32]; Ok, je l'sais. « amener », c'est amener des choses, pis « emmener », c'est emmener des personnes peut-être. [équipe 8, séq. 35]; C'est « amenait ». C'pasque m'semble d'habitude y'a une règle, là. Tsé, c'est comme... « embarquer ». En tout cas. J'me rappelle pas, c'est loin là, mais... [équipe 3, séq. 16]).

Par ailleurs, nous avons été surpris d'observer que les mêmes fausses règles ont été mentionnées par des équipes différentes. Par exemple, les équipes 3 et 8 ont tenté d'expliquer l'impossibilité de laisser le pronom *LUI* devant le verbe *INFORMER* dans **Elle lui informe ses déplacements* par une règle morphophonologique qui ne s'applique pas dans ce contexte (Ben, « lui », c'est pasque le mot « informe » commence avec un « l », faque tu peux pas dire « lui ». [équipe 8, séq. 44]; Ah! Pasque y'a deux « l » de suite... C'tu ça? Y'a deux voyelles... Donc, c'pour ça qu'on mettrait le « L » apostrophe. [équipe 3, séq. 25]). Deux équipes ont aussi invoqué de fausses règles pour justifier que la cooccurrence de *je* et de *pour ma part* constitue un pléonasme dans la phrase *Je crois pour ma part que d'autres formes de discipline sont aujourd'hui apparues* (Juste « je crois ». Pasque si tu dis « pour ma part », t'as pus de verbe, là. [équipe 5, séq. 3]; Y'a vraiment une erreur de syntaxe sur le fait que tu peux pas mettre « ma », pis « je », pis « personnellement », pis « je »... [équipe 1, séq. 30]).

Finalement, certaines « règles » étaient complètement bizarres et semblaient presque inventées de façon ad hoc pour expliquer un changement apporté par les sujets ([après avoir enlevé un *DONC* en début de phrase] Pasque j'trouve pas qu'on s'appuie sur grand-chose. « Donc », j'trouve que c'est très affirmatif. C'est plus, là, après cinquante pages, que t'affirmes : « Donc... » [équipe 7, séq. 19]; [pour justifier d'avoir remplacer « être porté de manquer » par « être porté à manquer »] Parce que c't'une indépendante et « de... que » ne va jamais ensemble, alors que « à... que », comme liaison, ça va ensemble. Mais j'pourrais pas te dire la règle exacte. [équipe 4, séq. 6]).

Les cas de référence à une autorité pour justifier un changement s'appuyaient sur des informations apprises à l'école (Moi, j'ai appris ça dans mes cours de diction au cégep. [équipe 7, séq. 23]; Tsé, c'est clair que depuis, m'semble, qu'on est au primaire on dit « à cause que », là, ça s'dit pas ! Faque ça serait la raison, là, pour laquelle j'dirais qu'ça s'dit pas. [équipe 2, séq. 32]), lues dans les médias (Ben, je l'sais pas si y'avait une règle, mais me semble que... dans la Presse, dans la chronique, j'ai lu quelque chose sur le mot « davantage » [équipe 2, séq. 4]; Je l'sais pus, là, où j'avais vu ça, mais y disaient la vraie formulation c'est « quant à moi ». [équipe 2, séq. 29]) ou provenant de membres de la famille (Ma grand-mère me l'a tellement dit souvent. « apporter », au lieu d' « emmener ». [équipe 7, séq. 20]). Certains commentaires de ce type laissent supposer (ou du moins espérer!) que le sujet a déformé les propos de la source d'autorité (Y'a vraiment une erreur de syntaxe sur le fait que tu peux pas mettre « ma », pis « je », pis « personnellement », pis « je »... [...] Y'a un prof qu'y a dit ça. [équipe 1, séq. 30]).

Une autre stratégie intéressante, bien que très peu utilisée, est l'**apport d'évidences linguistiques** pour prouver certains emplois d'unités lexicales. Deux exemples sympathiques ont été recueillis dans notre corpus : Ça s'dit pas « emporter »... Ben, ça

doit s'dire, là. « Autant en emporte le vent »... [équipe 7, séq. 20]; Ça existe « emmener »? [...] Ben oui, ben « emmène-moi »... [cherche dans le dictionnaire] Ah, c'est ça. « prendre avec soi quelqu'un ». Hey, je l'ai trouvé! J'suis fier de moi. C't'à cause de la chanson de Charles Aznavour que j'l'ai trouvé! [équipe 8, séq. 35]. En mentionnant ces titres connus, les sujets ont réussi à convaincre leur collègue de l'existence d'unités lexicales.

Nous terminerons ce tour d'horizon des types de commentaires les plus intéressants sur une touche humoristique. Voici un commentaire émis par un sujet venant de remplacer ENTHOUSIASTE par ENTHOUSIASME dans la phrase *L'enthousiaste des enfants pour leur nouveau compagnon a beaucoup diminué* : L'enthousiasme. C'est « me ». C'est la règle qu'on a appris au secondaire, mais j'pourrais pas te l'expliquer, c'est comme un automatique. [équipe 7, séq. 34].

2.2.3 LES STRATÉGIES DÉPLOYÉES EN FONCTION DES TYPES DE PROBLÈMES

La présente section s'intéressera aux principales stratégies déployées pour chacun des 13 types de problèmes lexicaux présents dans notre test. Nous tenterons de voir si les stratégies réellement mises en œuvre pour un type de problème particulier correspondent à celles que nous aurions pu prévoir et si elles sont adaptées au problème traité. Le tableau XV de la section V-1.2.3 présente les stratégies dominantes pour chaque type de problème. Avant d'entamer notre discussion à propos de stratégies observées, nous désirons rappeler que, comme pour le repérage des problèmes lexicaux (voir V-2.1.1), il se peut que le problème lui-même, plutôt que le type de problème, ait une influence sur les stratégies de résolution privilégiées. Par conséquent, devant d'autres problèmes des mêmes types que ceux apparaissant dans notre test, il est possible que les sujets aient opté pour d'autres stratégies.

À première vue, l'analyse du tableau XV (V-1.2.3) nous indique une grande variété de stratégies pour chaque type de problèmes. En effet, les faibles pourcentages des stratégies les plus fréquentes révèlent un nombre élevé de stratégies : si la stratégie dominante ne représente que 20 % du total des stratégies déployées, c'est que de nombreuses autres ont été mises en œuvre. Nous aurions pu nous attendre à ce que les problèmes les mieux résolus présentent clairement des stratégies de résolution « dominantes ». Pourtant, dans la plupart des cas, la différence entre les taux d'utilisation de la stratégie la plus fréquente et de la deuxième plus fréquente est peu marquée. Ceci nous indique que rares sont les types de problèmes clairement associés

à une stratégie particulière. Voyons maintenant ce qui ressort de l'analyse des commentaires visant la résolution du problème associés à chaque type d'erreurs.

Le non-respect des propriétés inhérentes de l'unité lexicale :

Les problèmes relatifs à cette dimension de l'unité lexicale constituent non seulement les mieux résolus, mais aussi ceux auxquels les stratégies de résolution déployées sont le mieux adaptées. En effet, dans une forte proportion (31,6 %), les sujets ont spontanément eu recours à une propriété de combinatoire grammaticale (Pis me semble « idole » c'est un nom féminin, non? [équipe 5, séq. 17]; « l'enthousiasme »... Pasque c'est le nom, et non pas un adjectif. [équipe 4, séq. 40]), ce qui constitue la stratégie idéale pour ce type de problème; elle indique que les sujets ont non seulement identifié la propriété de combinatoire à la source du dysfonctionnement, mais ont immédiatement eu accès à la solution en puisant dans leurs connaissances. La deuxième stratégie la plus fréquente pour ce type de problème est le recours au dictionnaire (17,5 %) et va de pair avec la première; les sujets ont encore une fois déterminé qu'il y a un problème relatif à la combinatoire grammaticale, mais doivent se référer au dictionnaire pour vérifier (« idole ». Ben oui. C'est un nom féminin. [équipe 2, séq. 24]; [vérifie dans le dictionnaire si LEÇON est invariable] C'est ça. Ben « leçons et devoirs », eux autres, ici, y l'mettent au pluriel. [équipe 2, séq. 11]).

L'usage de mauvaises prépositions :

Les problèmes de cette nature ont aussi été résolus correctement lorsqu'ils ont été traités. Par contre, les stratégies mises en œuvre révèlent un manque de ressources des sujets pour résoudre de tels problèmes avec des stratégies suffisantes. La stratégie la plus fréquente (17,1 % des cas) a été la proposition catégorique, c'est-à-dire que les sujets ont simplement corrigé le problème sans aucune explication ni justification (Ben là, « de manquer », là. « En effet, ceux-ci sont plutôt portés à manquer ». [équipe 8, séq. 11]). Une autre des stratégies dominantes pour ce type de problème est le recours au jugement du locuteur (14,6 %), ce qui signifie que les sujets avaient leur propre intuition comme seule justification de la correction apportée (Ben, j'sais pas, ça s'dit mieux, là. J'connais pas la règle, là, mais me semble... [équipe 3, séq. 3]; Moi, j'l'aurais sûrement pas écrit comme ça. [équipe 1, séq. 26]). Deux autres stratégies ont été utilisées aussi souvent que le jugement du locuteur (14,6 %) : la référence à des propriétés de combinatoire grammaticale et la consultation de dictionnaires (C'comme « porté à » et non « porté de ». [équipe 6, séq. 10] ; [cherche dans le dictionnaire] Être porté à! Ouais, c'est

ça. C'est c'qu'on a dit. Avoir l'habitude de. [équipe 2, séq. 8]). La correction non justifiée et le recours au jugement du locuteur ne constituent pas de mauvaises stratégies en soi et s'avèrent efficaces lorsque la personne est familière avec le problème auquel elle est confrontée, mais il peut paraître alarmant qu'elles constituent les stratégies les plus fréquentes. Nous croyons par contre que ceci est peut-être accidentel et qu'un problème de préposition régie différent aurait peut-être fait varier les stratégies privilégiées. En effet, il faut se rappeler que le problème de préposition régie ayant occasionné le plus de verbalisations est *être porté de* (plutôt que *être porté à*), problème à propos duquel l'ensemble des sujets avait une très forte intuition. Une seule équipe a traité l'autre problème de ce type, qui semble moins évident : *avoir de l'intérêt envers qqch.* Aucun des deux sujets n'étant parfaitement convaincu du choix de la préposition, ils ont eu recours aux deux autres stratégies mentionnées plus haut (Me semble on a un inter... un intérêt pour quelque chose et pas un intérêt envers quelque chose. [prend le dictionnaire] J'vas checker si y'a un exemple, ok? [...] « Son intérêt pour le sport » ! [équipe 1, séq. 26]).

La périphrase :

La périphrase que nous avons incluse dans notre test – *Un grand parc longeait ma rue de résidence* – nous semblait problématique parce que lourde et peu idiomatique. Nous avions anticipé que les sujets proposeraient la phrase *Un grand parc longeait ma rue* comme solution et justifieraient ce changement par des critères esthétiques. À notre étonnement, les sujets ont presque tous proposé la correction *Un grand parc longeait la rue de ma résidence*, qu'ils ont justifiée par des raisons sémantiques. Une seule stratégie se démarquait donc énormément des autres pour ce problème particulier : l'analyse sémantique de la phrase (Ça donne l'impression que c'est ma rue... toute la rue lui appartient [...] C'est comme si a l'habitait sur la rue. Faque ça peut avoir cette connotation-là, là. La rue de MA résidence. [équipe 3, séq. 24] ; C'pas « sa rue », c'est sa résidence qui est sur sa rue. [équipe 5, séq. 33] ; C'est sûr que sa résidence est sur sa rue, non? [équipe 8, séq. 43]). Une périphrase différente nous aurait peut-être permis d'observer d'autres stratégies.

Le non-respect du caractère figé d'une locution :

Deux types de stratégies se démarquent dans la résolution des problèmes de non-respect du caractère figé d'une locution, chacune représentant 12,9 % de l'ensemble des stratégies déployées pour ce type de problème (ce qui indique une grande variété).

Dans certains cas, les sujets se sont contentés de faire référence à l'expression utilisée en la mentionnant sous sa forme correcte, ce qui constitue une stratégie tout à fait pertinente pour de tels problèmes (L'expression c'est « prendre en main ». [équipe 2, séq. 36] ; Parce que l'expression c'est « le juste milieu ». [équipe 8, séq. 33]). L'autre stratégie privilégiée pour ces problèmes est le recours au sens; les sujets tentent de justifier la forme de l'expression figée par des raisons sémantiques (J'avais dit « leur permet de prendre en mains la responsabilité de leurs études ». Pasque « prendre entre leurs mains », c'est prendre entre leurs mains [montre ses mains]. [équipe 8, séq. 42] ; Ouais. « entre leurs mains », on dirait que c'est plus physique. [équipe 6, séq. 33] ; [le juste et bon milieu] Le juste milieu. Le juste milieu, c'est le bon milieu. Ça veut dire la même affaire. [équipe 7, séq. 17]). Dans ces cas-ci, les justifications sémantiques n'étaient pas trop douteuses, mais ce ne serait pas le cas avec toutes les locutions. Par exemple, comment expliquer, d'un point de vue sémantique, qu'on ne peut pas dire **se mettre quelque chose sous les dents*? Nous pouvons donc affirmer que le recours au sens n'est pas une stratégie très appropriée aux problèmes relatifs au caractère figé des locutions. La consultation d'un dictionnaire aurait été plus efficace pour vérifier la forme des expressions idiomatiques; pourtant, une seule équipe y a eu recours (« en main », c'est... singulier. Euh, non! C'est ici au pluriel. [...] Mais encore là, y'a pas de règles. « remettre une lettre en main (s) propre (s) », donc avec les « S » entre parenthèses, donc j'pense que c'est peut-être au choix. Dans l'fond, les deux c'est correct. [équipe 5, séq. 32]).

Les erreurs d'orthographe lexicale :

Les deux stratégies de résolution déployées devant les problèmes de forme de notre test présentent la même fréquence (23,3 %) et sont tout à fait appropriées à ce type de problème. Dans le cas le plus simple, les sujets connaissaient la forme correcte et ne faisaient que la mentionner (« davantage », ça prend pas un apostrophe. [équipe 3, séq. 7]). S'ils l'ignoraient, le recours au dictionnaire leur a permis de trouver la réponse à leur problème (« blanchiment », ça s'écrit pas avec un « E »? [trouve « blanchiment »] Non, r'gard. [équipe 7, séq. 22]). Nous pouvons donc affirmer que les sujets disposent de stratégies efficaces pour régler des problèmes relatifs à la forme.

Les problèmes de registres de langue :

Beaucoup des commentaires (19,3 %) concernant les problèmes de registres de langue véhiculaient une référence à la norme fonctionnelle (J'trouve que le terme « déprimer »... « ils déprimeraient moins », c'est comme si la personne écrit comme a parle, pis... [équipe 3, séq. 9]; J'pas sûr que, genre, quand t'écoutes TV5, la madame a dit « question de salir sa

réputation »... C'est sûr que si t'écoute Jean-Luc Mongrain, y va dire ça, là, mais... [équipe 1, séq. 39]; C'est comme... trop familier. Un ti peu. [équipe 6, séq. 32]). Les cas où les sujets ont tenté de régler des problèmes de cette nature en ayant recours au dictionnaire ont été peu fructueux, puisqu'ils s'attardaient sur la définition de l'unité lexicale sans lire les informations relatives aux niveaux de langue (Oui... « être démoralisé. Ils dépriment devant les innombrables travaux à remettre ». [équipe 4, séq. 18]). Pourtant, le recours au dictionnaire devrait être une stratégie efficace pour ce type de problème; il s'agirait que les étudiants prennent conscience du fait que les unités lexicales appartenant à d'autres registres de langue que le registre standard sont aussi décrites et que, par conséquent, ce n'est pas parce qu'un « mot » est dans le dictionnaire qu'il appartient nécessairement au français standard. La troisième stratégie qui se démarque pour les problèmes de registres de langue est le recours au jugement du locuteur (7,9 %), mais cette stratégie a davantage servi à repérer le problème qu'à le solutionner (J'aime pas ben, ben ça d'« être déprimés de ». J'pense j'ai jamais dit ça « j'déprime de » [équipe 2, séq. 26]; Non. Y'a pas d'rai... Ben, moi c'est juste comme ça à l'oreille, là. [équipe 2, séq. 42]).

Le pléonasme :

Près du tiers des commentaires visant la résolution du pléonasme de notre test comportaient une analyse sémantique de la phrase dans laquelle il apparaissait, ce qui est compréhensible, puisqu'un pléonasme naît de la cooccurrence de deux termes (Quin. « plus les générations deviennent », pis là « DAVANTAGE permissives ». Y'a déjà le « plus » en avant. [équipe 1, séq. 2]; Ben, c'pasque tu répètes deux fois la même chose. « pour ma part » pis « je », tu répètes tsé. Tsé, si tu dis « je crois », ben c'est sûr que c'est pour ta part. [équipe 8, séq. 2]; « Certains disent qu'avec le temps, les générations sont devenues plus permissives ». C'est toujours dire « plus, plus, plus » [équipe 7, séq. 1]). Dans de plus faibles proportions (9,4 %), les sujets ont eu recours au sens d'une des unités lexicales constituant le pléonasme (Le « davantage » y'est utilisé ici pour dire « plus » dans l'fond. [équipe 7, séq. 1]; Parce que « permissives », c'est être plus tolérant, donc encore là, le « davantage » devient de trop. [équipe 8, séq. 6]) ou à leur propre jugement de locuteur lorsqu'ils n'arrivaient pas à justifier les changements effectués (Mais j'sais pas, ça sonne bizarre. C'est juste à l'oreille peut-être, là, j'trouve... [équipe 2, séq. 4]; J'sais pas, j'aime pas ça. [équipe 3, séq. 15]). Nous pouvons donc dire que les stratégies privilégiées dans la résolution des pléonasmes sont bien adaptées au type de problèmes.

Les problèmes de sens :

Les problèmes de sens étaient nombreux dans notre test; nous avons donc pu analyser plusieurs commentaires concernant ce type de problèmes. Évidemment, les stratégies de résolution employées par les sujets étaient extrêmement variées. Il ressort quand même quatre stratégies principales de notre analyse, qui semblent très bien convenir à la résolution de tels problèmes. La plus fréquente (20,9 %) est le recours au sens des unités lexicales en jeu pour en proposer une définition ou pour mettre en évidence une composante de sens (Ben oui, pasque « affecter », c'est « avoir un effet sur quelque chose ». [équipe 4, séq. 23]; On dirait qu'on explore un lieu, mais pas un objet. [équipe 5, séq. 37]; « Pour en terminer avec la discipline », c'est clair qu'ça s'dit pas, là. [...] On dirait qu'y va y casser la face, je l'sais pas! [équipe 1, séq. 20]). La seconde stratégie la plus fréquente (17,7 %) pour les problèmes de sens est le recours au dictionnaire (Laxisme, nom masculin. [...] « tendance marquée à la conciliation, à la tolérance ». J'imagine que c'correct, on connaît juste pas l'mot. [équipe 1, séq. 23]; [cherche « emmener » dans le dictionnaire] « è-mener ». « è-mener. Mener avec soi quelqu'un, un animal. Prendre avec soi en partant. » [équipe 4, séq. 26]). Une autre stratégie fréquente (9,4 %) consistait à analyser sémantiquement la phrase traitée dans le but de déterminer si l'unité lexicale utilisée est compatible avec le contexte (C'pasque... dans l'fond, c'qui voulait dire, c'est qu'les deux groupes « collaborent », et non pas « corroborent ». Pas' c'est pas... c'pas « appuyer », là. [équipe 2, séq. 14]; [... les professeurs doivent parfois imposer des conséquences aux élèves.] Ben, des conséquences, encore là... à quoi ? Tsé, pasque la phrase d'avant, c'est les profs qui doivent utiliser leur autorité pour faire comprendre l'importance des études. Mais, les conséquences, ça pas rapport. [...] Tsé, tu donnes pas des conséquences. [équipe 5, séq. 10]). Une quatrième stratégie de résolution a été utilisée assez fréquemment (6,7 %) par les sujets : la comparaison de deux solutions. Les problèmes de sens sont les seuls pour lesquels cette stratégie a été employée de façon importante. Par exemple, les sujets de l'équipe 5 désiraient trouver un verbe pour remplacer ATTAQUER dans *la paresse attaque les jeunes* et hésitaient entre AFFLIGER et AFFECTER : [lit l'article d'AFFLIGER] « toucher de façon pénible ». C'pasque ça touche pas de façon pénible. Tsé, les jeunes trouvent pas ça pénible d'être paresseux. [...] Ben, « affecter », c'est peut-être moins euh... moins gros comme... comme une pression que ça fait là. C'est plus général comme mot. « affliger », ça l'air de mettre un gros poids, là. [remplace « affliger » par « affecter »] [équipe 5, séq. 4]. La même équipe a employée cette stratégie une seconde fois avec l'impropriété CONSÉQUENCE (au sens de 'punition') : « Les professeurs doivent parfois imposer des... sanctions ». « sanctions » ? Ouais, mais « sanction » c't'un synonyme, là. Ben, au moins ça se dit plus que « imposer des conséquences ». [regarde dans le dictionnaire] La j'viens de

chercher « sanction », pis j'trouve que ça va bien... que ça remplacerait bien « conséquence » ». [équipe 5, séq. 10].

Les anglicismes :

Les commentaires concernant les anglicismes ont été plutôt rares dans le corpus, puisque les problèmes de cette nature n'ont été discutés que par un petit nombre d'équipes. Une seule stratégie se démarque dans la résolution de tels problèmes : le recours au dictionnaire (J'aime le Multidictionnaire! « mettre l'emphase sur : anglicisme au sens de mettre l'accent sur, mettre l'insistance sur, insister sur, faire ressortir. » [équipe 5, séq. 12]). Selon nous, cette stratégie est la seule valable dans le cas des anglicismes, puisque la décision de condamner certains usages pourtant très répandus est somme toute assez arbitraire⁴⁰; en cas de doute, la vérification au dictionnaire constitue la seule solution possible. Certains sujets ont tout de même justifié leur doute en établissant un lien avec une unité lexicale de l'anglais, mais cette technique était plutôt marginale (« opportunity ». En anglais, c'est très utilisé. [équipe 4, séq. 12]).

Les problèmes de collocations :

Nous avons vu à la section V-1.1.2 que les séquences portant sur les problèmes de collocations présentent des stratégies insuffisantes dans 52,6 % des cas. L'analyse des commentaires relatifs à ces problèmes nous a permis de mieux comprendre pourquoi. La stratégie la plus souvent mise en œuvre dans la résolution de mauvaises collocations est le recours aux ouvrages de référence, ce qui peut sembler adéquat. Par contre, comme nous l'avons mentionné au point V-2.2.2, les étudiants ne savent pas comment chercher des informations relatives aux collocations dans le dictionnaire (J'essaie d'avoir où est-ce qu'on pourrait trouver ça. [prend le dictionnaire] Ben, j pense pas qui nous l'indiquent, mais... [équipe 2, séq. 40]) et tentent la plupart du temps de les trouver dans les articles des collocatifs ([les parents devraient avoir plus de discipline avec leurs enfants] Ben, j'vas faire quelque chose... J'vas aller voir « avoir ». Qu'est-ce ça veut dire « avoir »? [équipe 7, séq. 31]). Une autre stratégie fréquemment employée pour les problèmes de collocation est le recours au sens (10,1 %), même si cette façon d'aborder ce type de problème n'est vraiment pas efficace (« appliquer un plus grand respect de soi » ? « appliquer », c'est aller chercher trop loin. [prend le dictionnaire]

⁴⁰ Nous n'avons qu'à penser au Multidictionnaire qui condamne l'expression *prendre une marche* qui est considérée comme un calque de *to take a walk*, alors que *PRENDRE* semble être le collocatif le plus naturel pour exprimer ce sens auprès de *MARCHE*. Cet ouvrage propose plutôt l'expression *faire une marche*, qui semble bizarre pour la très grande majorité des locuteurs québécois que nous avons interrogés.

« appliquer : apposer. Appliquer une couche de peinture. Diriger son attention : appliquer son esprit à... Donner toute son attention. Elle s'est appliquée à étudier la chimie. » Faque dans l'fond, faut pas dire « appliquer » ici. [équipe 7, séq. 11]). En effet, il semble tout à fait correct de dire *appliquer des principes/valeurs comme le respect* et le problème découle de la cooccurrence des unités lexicales RESPECT et APPLIQUER, ce dont on ne peut rendre compte par des raisons sémantiques. Finalement, une autre stratégie parmi les plus employées pour les problèmes de collocations est la proposition catégorique (12,7 %), ce qui signifie que les sujets ont corrigé le problème en ne s'appuyant sur aucune justification (Moi, j'dirais plus, tsé, « d'apprendre à se respecter et à respecter les autres ». [équipe 3, séq. 5]; Peut-être « faire preuve d'autorité face aux élèves ». [équipe 1, séq. 10]). Ceci renforce donc l'hypothèse selon laquelle les étudiants possèdent de fortes intuitions concernant les collocations, mais ignorent que les ouvrages de référence contiennent certaines informations relatives à cette dimension de l'unité lexicale. Par ailleurs, comme nous savons que la collocation est un phénomène très arbitraire et que le collocatif tend à perdre son sens au sein d'une telle expression (voir II-1.5.3), nous comprenons pourquoi le fait d'aborder ce type de problèmes d'un point de vue sémantique ne peut que conduire à une impasse.

Le non-respect du mode verbal imposé par l'unité lexicale :

Deux stratégies dominantes ont été utilisées dans les mêmes proportions (16,7 %) dans la résolution du problème de non-respect du mode verbal imposé par l'unité lexicale BIEN QUE. La première stratégie a consisté à se fier à son intuition de locuteur pour changer le mode indicatif pour le subjonctif (*Est-ce qu'y a une raison pourquoi qu'c'est l/subjonctif?* Euh... À l'oreille, là, mais... [équipe 2, séq. 6]; Ben... Moi j'l'aurais écrit d'même. [équipe 6, séq.6]). L'autre stratégie souvent mise en œuvre est l'analyse grammaticale de la phrase dans le but de trouver l'élément qui demande le subjonctif; dans aucun cas la locution n'a été mentionnée (Ne soit... C'pas à cause du « que », là ? « Que... je »... [équipe 1, séq. 4]; À cause du « que » qui introduit le subjonctif. [équipe 6, séq.6]). Pour ce type de problème particulier, les stratégies idéales auraient été la référence directe à la propriété de combinatoire grammaticale de l'unité lexicale BIEN QUE, ou alors la consultation du dictionnaire, qui mentionne le mode imposé par la locution. Par contre, comme aucun sujet n'a perçu la locution comme un élément de la solution, il est compréhensible qu'aucun n'ait eu le réflexe de chercher la locution dans les ouvrages de référence. À notre avis, les étudiants ne sont pas conscients que certaines unités

lexicales imposent un mode verbal et encore moins du fait que des locutions ont des propriétés semblables.

Les problèmes de non-respect de la structure actancielle des unités lexicales :

Les problèmes de non-respect de la structure actancielle figurent parmi les plus mal résolus (56,3 %) et les séquences les concernant ont presque toujours été considérées insuffisantes (87,5 % des cas). En observant les stratégies mises en œuvre dans leur résolution, nous constatons d'ailleurs qu'elles sont extrêmement variées et très souvent peu adaptées au problème. La stratégie la plus fréquente ne représente que 13,3 % de l'ensemble des stratégies déployées; il s'agit du recours au dictionnaire. Par contre, nous ne croyons pas que cette observation soit réellement représentative des stratégies privilégiées pour de tels problèmes, puisque l'ensemble des recours au dictionnaire portaient sur la phrase **j'ai décidé de donner l'accès à mon frère* et que la totalité des sujets ayant opté pour cette stratégie s'interrogeaient en fait à propos de l'utilisation du verbe DONNER avec le nom ACCÈS (De donner l'accès? Permettre l'accès? Peux-tu checker dans l'dictionnaire? Dans « accès » ou dans « donner »? [équipe 1, séq. 34]; « avoir accès... Donner accès! » Ben oui! « permettre d'entrer ». Faque finalement... Donner l'accès. [...] Peut-être que nos premières impressions étaient fausses. [équipe 2, séq. 30]). Les deux autres stratégies employées, bien que moins fréquentes (9,6 %), convenaient davantage à la résolution du problème et concernaient majoritairement la phrase **elle lui informe ses déplacements*. La première est la référence à des propriétés de combinatoire grammaticale (« elle l'informe », c'est ça ? Pasque qu'on informe quelqu'un, on n'informe pas « à » ou « de » quelqu'un, là. [équipe 2, séq. 38]; Pasque « informer » [rires] vient avec le mot « de ». [équipe 8, séq. 44]) et la seconde l'analyse grammaticale de la phrase (C'est... ça prend un « L » apostrophe qui va être là pour euh... L'inspecteur. [équipe 2, séq. 38]; « de lui en donner », sauf que... on n'a pas besoin d'mettre « lui », pasque « à mon frère » y'est là. Tsé, c'est comme « je donne à elle » ou « je lui donne ». [équipe 6, séq. 27]).

L'utilisation d'un mauvais auxiliaire :

Le problème d'utilisation du mauvais auxiliaire dans **il avait même monté dans le mauvais autobus* n'a été bien résolu que dans la moitié des cas. Deux stratégies ressortent dans la résolution de ce problème. La première est la référence directe au fait que le verbe MONTER appelle l'auxiliaire ÊTRE dans ce contexte (Parce que c'est un verbe que au passé composé, il se conjugue avec « être », pas avec « avoir ». [équipe 4, séq. 30]; Mettons que tu conjugues le verbe « monter »... Tu vas pas dire « j'ai monté ». C'est « je suis monté ». [équipe 7, séq. 24]) et représente 14,3 % des stratégies déployées pour ce

problème. La deuxième stratégie dominante est, encore une fois, commune aux sujets qui n'ont pas traité correctement le problème présent dans cette phrase. En effet, quelques équipes croyaient que le problème était le choix du verbe MONTER et, par conséquent, le recours au sens constitue 11,9 % des commentaires relatifs à cette phrase (Ben, c'pas « monter », là. « embarquer »... Pasque « monter », c'est pas l'même sens. [équipe 3, séq. 20]; « embarquer », c'est un bateau ou un train. [équipe 6, séq. 30]). Toutefois, une équipe a tenté de justifier le choix de l'auxiliaire par des raisons sémantiques, ce qui ne fonctionne pas avec le verbe MONTER (Pasque « avoir », c'est quand t'as quelque chose. Tandis que le verbe « être »... Tu es. [équipe 7, séq. 24]).

CONCLUSION

Nos observations nous permettent donc d'affirmer que les stratégies de résolution privilégiées par les étudiants étaient, de façon générale, assez bien adaptées aux problèmes traités. Par contre, nous devons rappeler que les pourcentages de fréquence sont peu élevés, ce qui signifie que plusieurs stratégies peu adéquates ont été utilisées dans beaucoup de cas. Ceci démontre un certain tâtonnement dans la résolution de problèmes lexicaux et découle en grande partie du fait que les étudiants ne sont pas conscients de certains phénomènes lexicaux (prépositions régies, collocations, etc...) ni de la variété d'informations contenues dans les dictionnaires. Par ailleurs, bien que nous puissions nous réjouir du fait que les stratégies observées sont la plupart du temps appropriées aux problèmes traités et que ces derniers sont généralement bien résolus, il ne faut pas oublier qu'un bon nombre de problèmes n'ont pas été repérés et sont demeurés non résolus.

De plus, certaines stratégies très répandues sont peu sophistiquées et ne permettent que de régler les problèmes sans vraiment en comprendre la cause. Le recours au jugement de locuteur, par exemple, peut constituer une très bonne stratégie, puisque les intuitions exprimées s'avèrent correctes dans la plupart des cas. Par contre, cette stratégie devrait idéalement être combinée avec d'autres si l'étudiant voulait diagnostiquer le problème de façon plus approfondie.

VI. CONCLUSION

1. RAPPEL DES PRINCIPAUX RÉSULTATS

Nous voilà donc au terme de notre recherche sur la maîtrise lexicale des étudiants universitaires. Nous en rappelons brièvement les principaux résultats.

Le premier objectif visait la description des problèmes lexicaux rencontrés dans des textes produits par des étudiants universitaires. Il ressort entre autres de cette analyse que près du tiers des problèmes lexicaux sont relatifs à une transgression des propriétés de combinatoire des unités lexicales. Les autres types majeurs de problèmes lexicaux sont les erreurs d'orthographe lexicale, les impropriétés et l'utilisation d'unités lexicales appartenant à un registre de langue inapproprié. Nous soulignons que la grande majorité de ces problèmes auraient pu être évités si les étudiants avaient consulté le dictionnaire de façon efficace. La réalisation de ce premier objectif de recherche nous a aussi démontré qu'il est très difficile, d'un point de vue didactique, de s'en tenir à une classification structurale des problèmes lexicaux sans en considérer les sources possibles.

Notre deuxième objectif de recherche visait la description des stratégies de résolution de problèmes lexicaux des étudiants universitaires. Nous avons constaté que la moitié des problèmes lexicaux présents dans la première partie de notre test n'ont pas été repérés. Par contre, plus de 80 % des problèmes traités ont été résolus correctement et les stratégies de résolution privilégiées par les étudiants étaient, de façon générale, assez bien adaptées aux types de problèmes traités. De l'analyse de leurs verbalisations, nous pouvons conclure que les étudiants ont de la difficulté à expliquer certains phénomènes lexicaux (prépositions régies, collocations, etc.), bien qu'ils semblent en avoir une connaissance intuitive, et qu'ils utilisent mal le dictionnaire parce qu'ils ne sont pas familiers avec les informations qu'il contient.

En mettant en parallèle les résultats des deux volets de notre recherche, nous pouvons faire certaines constatations intéressantes.

Tout d'abord, les problèmes de sens, qui constituent près du quart des erreurs observées dans les copies d'étudiants, figurent pourtant parmi les mieux repérés au deuxième volet et présentent un taux de résolution assez élevé. Les stratégies déployées pour ce type de problèmes sont généralement appropriées et suffisantes. Pourquoi les productions présentent-elles autant de problèmes de sens si le volet de révision semble indiquer que les étudiants disposent des outils nécessaires à une résolution efficace de tels problèmes? À notre avis, cela s'explique par le fait que les étudiants doutent peu de leur maîtrise du sens des unités lexicales. En apprenant à

douter à bon escient et à faire la différence entre les unités lexicales qu'ils maîtrisent réellement et celles dont ils ont une connaissance approximative, ils pourraient éviter une bonne partie des problèmes de sens.

Ensuite, nous remarquons que les problèmes de régime (structure actancielle, prépositions régies et mode verbal) sont assez fréquents dans les copies d'étudiants (13 %). En observant les stratégies déployées pour ces types de problèmes, nous constatons certaines lacunes. En effet, les stratégies s'appuient surtout sur l'intuition et dénotent un manque de maîtrise de la notion de régime. En étant plus conscients de la microsyntaxe régie par les unités lexicales, les étudiants seraient plus à même d'utiliser les ouvrages de référence pour vérifier certaines informations relatives à des unités lexicales particulières.

Finalement, la maîtrise de la norme fonctionnelle semble défailante, notamment en ce qui concerne la différence entre le code écrit et le langage parlé. En effet, près de 10 % des problèmes relevés dans notre corpus présentaient l'emploi d'unités lexicales incompatibles avec le registre de langue exigé par le contexte de rédaction scolaire. Pourtant, ces problèmes ont été résolus de façon satisfaisante dans le cadre du deuxième objectif de notre recherche. Par contre, nous constatons encore une fois que les stratégies déployées reposent sur l'intuition et que les recours au dictionnaire pour ce type de problèmes se sont avérés infructueux parce que les étudiants ne savaient pas où chercher les informations pertinentes.

En bref, nous comprenons que, si les problèmes lexicaux sont présents dans les copies d'étudiants universitaires, c'est qu'ils ne savent pas « douter stratégiquement » parce qu'ils ne connaissent pas les principaux phénomènes en jeu dans l'usage du lexique et qu'ils manquent de ressources pour utiliser efficacement le dictionnaire qui s'avère un atout précieux dans la maîtrise des unités lexicales.

2. LES RECHERCHES FUTURES

Si une recherche qui s'intéresse à une dimension peu explorée procure bien sûr quelques réponses, elle fait surtout émerger une multitude de questions. Le simple fait d'effectuer des choix méthodologiques nous oblige à renoncer à l'exploration de certaines avenues pourtant intéressantes et à envisager notre problématique selon une perspective particulière. La présente recherche n'échappe pas à cette situation et les

limites de la zone explorée nous laissent avec de nombreuses questions, qui pourraient faire l'objet de recherches futures. Voici les principales :

- Les textes analysés dans le cadre de notre premier objectif de recherche provenaient tous d'étudiants universitaires francophones inscrits à la même faculté. Bien sûr, il serait possible de prendre en considération d'autres échantillons. Premièrement, il pourrait être intéressant de disposer de textes d'étudiants universitaires issus de différentes facultés afin de s'assurer que notre échantillon était effectivement représentatif de la population universitaire en général. Deuxièmement, il faudrait décrire les problèmes lexicaux des élèves du primaire et du secondaire afin de vérifier s'il y a une évolution dans les types de problèmes rencontrés. Ceci serait particulièrement important si l'on considère que des interventions sont davantage possibles à ces ordres d'enseignement qu'à l'université. Troisièmement, il faudrait analyser les problèmes lexicaux d'étudiants n'ayant pas le français comme langue première pour constater si les problèmes rencontrés sont les mêmes et s'ils présentent de grandes différences en ce qui concerne la fréquence de chaque type. Une autre variable qui pourrait être considérée est le niveau de maîtrise du code écrit des sujets dont les rédactions sont analysées. Finalement, il serait très intéressant de transposer notre typologie des problèmes lexicaux à une autre langue, par exemple l'anglais, pour vérifier si elle s'applique aussi et quels sont les changements à y apporter pour l'adapter; une telle transposition contribuerait à la validation de notre typologie.
- Nous avons constaté la présence d'une moyenne de 3,8 problèmes lexicaux par copie dans notre corpus. Il serait intéressant de considérer, lors d'analyses futures, les erreurs relatives à toutes les dimensions du code (orthographe grammaticale, syntaxe, cohérence textuelle, etc.). De cette façon, il serait possible d'établir clairement la proportion d'erreurs lexicales par rapport à l'ensemble des erreurs.
- Étant donné le nombre limité de sujets (16), nous pouvons considérer que notre deuxième objectif de recherche était plutôt exploratoire. Observer les stratégies de résolution de problèmes lexicaux d'un plus grand nombre d'étudiants contribuerait à cerner mieux les processus en jeu. Il serait aussi important d'augmenter le nombre de problèmes traités afin de pouvoir établir davantage de tendances entre les types de problèmes, leur repérage, et les stratégies déployées dans leur résolution. Les variables mentionnées plus haut pourraient aussi être considérées (force en français écrit, fait d'avoir le français comme langue première ou seconde, domaine et niveau d'études, etc.).
- Nous avons cité quelques auteurs, dans notre problématique, qui affirmaient que le lexique est l'objet d'un enseignement plutôt sporadique et souvent inefficace (Simard, 1994; Lefrançois *et al.*, 2005; Dreyfus, 2004). Il serait intéressant de voir de quelle façon les enseignants perçoivent et enseignent le lexique. Quelles sont selon eux les connaissances nécessaires à la maîtrise d'une unité lexicale? Que considèrent-ils comme un problème lexical? Disposent-ils des ressources et outils suffisants pour trouver des réponses à leurs questions relatives au lexique? Quelles sont leurs propres stratégies de résolution de problèmes lexicaux? Comment expliquent-ils ce type d'erreurs à leurs élèves/étudiants?

3. LINGUISTIQUE ET DIDACTIQUE

Avant de conclure notre mémoire, il est important de mentionner un problème auquel nous nous sommes heurté tout au long de cette recherche : la difficulté d'allier linguistique et didactique. En effet, nous avons sans cesse vécu des tiraillements entre une vision plutôt descriptive de la langue, propre à la linguistique, et une approche normative associée à la didactique. Ceci découle bien sûr du fait que la norme lexicale est, comme nous l'avons vu, plutôt difficile à définir. Par conséquent, nombreux sont les cas à propos desquels le degré d'acceptabilité variait énormément d'une personne consultée à une autre.

Nous avons tenté de rendre compte de ce fait en introduisant un axe d'analyse « statut du problème » qui visait à départager les erreurs réelles des simples maladroresses lexicales. Par contre, force est de constater que cette bipartition des problèmes est bien grossière et ne rend que très peu compte de la difficulté de trancher à propos du statut d'un problème lexical. Il semble en effet qu'il existe tout un spectre de nuances entre une « erreur » et une « maladresse ». De plus, certains problèmes que nos choix méthodologiques nous ont poussé à qualifier d'erreurs ne nous paraissaient pas aussi facilement condamnables que d'autres, notamment parce qu'ils sont extrêmement fréquents et témoignent selon nous d'une évolution normale de la langue qui laisse présager qu'ils constitueront peut-être la norme d'ici quelques temps. Nous pensons notamment à des cas comme l'emploi du subjonctif après le verbe ESPÉRER (en *espérant que* cette jeune femme *ait* une meilleure vie [sujet G-21]) ou le soi-disant anglicisme *tomber en amour* qui constitue à nos yeux un québécoïsme tout à fait légitime. Peut-être serait-il utile de considérer le degré d'intégration du problème au code linguistique lorsque nous tentons de statuer sur sa « gravité ».

Nous avons souligné la difficulté de déterminer si un problème constitue réellement une erreur ou simplement un emploi un peu maladroit d'une unité lexicale, mais le même flou existe entre ce que nous appelons un problème lexical et ce que nous considérons comme un usage acceptable d'une unité lexicale. En effet, il se peut que les problèmes que nous avons relevés lors de notre analyse de corpus ne soient pas considérés problématiques par d'autres locuteurs. De la même façon, certains usages que nous n'avons pas retenus parce qu'ils nous semblaient corrects auraient peut-être été condamnés par un autre correcteur.

Nous constatons donc que la question de la norme lexicale est un problème complexe qui peut avoir plusieurs implications pour les enseignants. En effet, ceux-ci se trouvent

dans l'obligation de prendre position par rapport à ces problèmes chaque fois qu'ils corrigent une copie. Doivent-ils se fier exclusivement aux ouvrages de référence, même lorsque ceux-ci vont à l'encontre de leurs propres intuitions? Peuvent-ils pénaliser un étudiant pour un problème qui relève exclusivement des normes descriptive ou fonctionnelle? S'ils choisissent de ne pas considérer certains usages douteux faute de pouvoir prouver qu'ils sont problématiques, doivent-ils tout de même les mentionner et en suggérer d'autres qui sont plus idiomatiques, naturels ou acceptables?

Nous tenons donc à souligner de nouveau que les problèmes lexicaux présents dans ce mémoire ont été retenus en vertu de certains choix méthodologiques et parce que nous travaillons dans une perspective plutôt normative que nous pensons typique du monde de l'enseignement. Par contre, nous sommes conscient que, d'un point de vue linguistique, une telle approche puisse sembler extrême et, à la limite, aller à l'encontre de toute liberté en ce qui concerne le lexique, qui est pourtant la dimension de la langue propice à la plus grande créativité. Nous voulons donc rassurer nos lecteurs en spécifiant que les problèmes analysés ici ne nous semblent pas tous condamnables, mais plutôt marqués par rapport à l'usage standard; par conséquent, ils n'auraient pas forcément été pénalisés dans un cadre scolaire. D'ailleurs, nous comptons bien continuer à « prendre des marches » et à « tomber en amour », même si ces plaisirs sont apparemment réservés à nos amis anglophones!

4. LES CONCLUSIONS

Quelles conclusions pouvons-nous donc tirer de toutes les informations présentées dans ce mémoire? Comme cette recherche en didactique du français a été menée en codirection avec le Département de linguistique et de traduction, nous présenterons d'abord quelques remarques relatives à la linguistique avant de présenter les pistes didactiques qui émergent de notre travail.

4.1 LES PISTES LINGUISTIQUES

À notre connaissance, la théorie Sens-Texte a servi d'assise à peu de travaux en didactique, notamment dans une perspective d'analyse d'erreurs. Nous croyons avoir démontré, par l'élaboration de notre typologie des problèmes lexicaux, que ce cadre théorique – à tout le moins la lexicologie explicative et combinatoire – permet une catégorisation fine et élégante de ce type de problèmes.

Bien sûr, nous avons constaté que les classes d'erreurs que nous avons anticipées d'un point de vue théorique ne correspondent pas toujours de façon exacte aux problèmes réels et qu'elles se superposent dans certains cas. Cette observation est peu surprenante si nous tenons compte du fait que la théorie Sens-Texte, à l'instar des autres théories linguistiques, vise plutôt la description d'une langue idéale. Polguère (à paraître) mentionne d'ailleurs ce problème relativement à l'analyse des greffes collocationnelles :

« [...] il nous faudra définir des notions qui n'appartiennent pas au cheptel notionnel de la théorie Sens-Texte « classique ». En effet, cette dernière se concentre exclusivement sur la modélisation de la langue et exclut la modélisation des phénomènes de parole de son champ d'étude. Nous sommes donc obligé d'ajouter l'embryon d'un module à la théorie Sens-Texte, module qui nous paraît nécessaire si l'on veut pouvoir passer de la description idéalisée du code linguistique à la modélisation de son utilisation » (Polguère, à paraître : 2)

Néanmoins, nous pensons que l'appareillage conceptuel de cette approche linguistique nous a permis de définir assez clairement la notion d'erreur lexicale et de délimiter davantage les frontières floues entre le lexique et la syntaxe. D'ailleurs, nous croyons que la didactique des langues, en particulier l'enseignement du lexique, gagnerait à s'inspirer davantage de la théorie Sens-Texte, qui considère le sens comme le moteur de la langue. À partir de l'étude de concepts plus intuitifs comme le sens et les « mots », les étudiants seraient peut-être plus à même de comprendre les mécanismes qui régissent le fonctionnement de la langue. Par exemple, le concept de structure actancielle – que l'on peut comprendre de façon assez intuitive – peut servir de point de départ à une explication de la notion de *transitivité*, qui semble poser problème à beaucoup d'étudiants. Ils pourraient ainsi revenir à cette notion simple pour mieux saisir des règles grammaticales plus complexes, comme l'accord des participes passés. À l'inverse, nous sommes convaincu que la théorie Sens-Texte retirerait beaucoup à s'intéresser davantage à l'étude de la parole. Nous pensons que l'étude de phénomènes

marginaux comme les erreurs lexicales pourraient permettre d'améliorer et de valider les outils conceptuels de la théorie, puisqu'elles constituent selon nous des indices de la façon dont s'articule le système linguistique dans la tête des locuteurs qui commettent ces erreurs.

4.2 LES PISTES DIDACTIQUES

Ce mémoire a démontré que les problèmes lexicaux sont bien présents dans les copies d'étudiants universitaires, que ces derniers ne les repèrent pas dans la moitié des cas et qu'il existe certaines lacunes dans les stratégies qu'ils utilisent pour résoudre les problèmes de nature lexicale. Ces constatations devraient nous convaincre de la nécessité de revoir certaines pratiques relatives à l'enseignement du lexique. Voici quelques pistes qui pourraient, à notre avis, favoriser un enseignement plus efficace du lexique, notamment à travers le traitement des problèmes lexicaux.

- Il faudrait que les enseignants aient une **conception claire de l'erreur lexicale**, conception qui prendrait en considération les trois dimensions de l'unité lexicale (forme, sens et propriétés de combinatoire). Une telle conceptualisation de la notion d'erreur lexicale leur permettrait de concevoir plus clairement les interactions entre le lexique et la grammaire, ce qui les aiderait à fournir aux étudiants des explications en accord avec le fonctionnement réel de la langue;
- Il est aussi essentiel que cette définition de l'erreur lexicale rende compte du fait que la norme lexicale est très floue et que **plusieurs types de normes sont à prendre en considération dans le diagnostic des problèmes lexicaux**; une approche purement prescriptive est impossible pour cette dimension du code écrit;
- Il faut que les étudiants eux-mêmes prennent conscience de la **quantité d'informations en jeu dans la maîtrise d'une unité lexicale**, particulièrement en ce qui concerne les propriétés de combinatoire;
- **L'analyse d'erreurs lexicales en classe peut constituer un exercice fort bénéfique**, puisqu'elle permet aux étudiants de constater la variété de propriétés inhérentes à l'unité lexicale, d'acquérir de nouvelles stratégies de résolution de problèmes lexicaux et d'apprendre à douter de certains aspects lexicaux qu'ils maîtrisent moins. Le fait de travailler ce type de problèmes avec un expert de la langue comme l'enseignant va mettre en lumière l'absence de justifications possibles dans certains cas. Ainsi, l'étudiant apprendra à se fier davantage à son intuition de locuteur pour certaines dimensions (collocations, caractère figé des locutions, etc.).

Comme nous l'avons souligné à quelques reprises, il est impossible de maîtriser, donc d'enseigner, l'ensemble du lexique d'une langue. Par conséquent, il sera plus efficace de fournir aux étudiants les outils nécessaires à l'acquisition de nouvelles unités

lexicales ou à l'approfondissement de la connaissance d'unités lexicales moins bien maîtrisées que d'exiger d'eux qu'ils apprennent un nombre limité de « mots nouveaux ». En d'autres termes, mieux vaut favoriser leur compétence lexicale que tenter d'augmenter leur maîtrise lexicale, ce qui devrait se faire naturellement au cours de leurs lectures et de leurs études. Pour ce faire, il faudrait :

- **enseigner les principaux phénomènes lexicaux** afin que les étudiants comprennent mieux de quelle façon le lexique impose ses caprices à la grammaire. Nous pensons notamment qu'une attention particulière devrait être portée aux notions de régime, de collocation et de registres de langue;
- **favoriser l'usage du dictionnaire**. Il est primordial d'enseigner aux étudiants à se servir d'ouvrages de référence, de les familiariser avec la structure des articles de dictionnaires et de leur faire comprendre la nature des informations qui y sont présentées. Leur faire comparer certains ouvrages leur permettrait aussi de développer des réflexes de recherche d'informations;
- **favoriser l'établissement de liens lexicaux** afin que les étudiants perçoivent le lexique comme un tout cohérent possédant une certaine structure et non comme un regroupement désordonné d'unités lexicales (à l'image des dictionnaires commerciaux qui présentent le lexique comme une suite linéaire d'unités lexicales n'ayant aucun lien entre elles sinon celui de « commencer par la même lettre »). La présentation et l'exploration de notions comme la synonymie, l'antonymie, la polysémie et l'homonymie contribueraient, entre autres, à une perception plus cohérente du lexique en mettant en lumière les relations et les oppositions qui servent à tisser ce réseau complexe d'unités lexicales.

Mais, plus que tout, il est important que les enseignants saisissent l'importance de la maîtrise lexicale dans la maîtrise du code en général et qu'ils communiquent la curiosité et l'enthousiasme qui doivent faire partie de l'apprentissage lexical. En permettant à leurs étudiants de mieux saisir le fonctionnement du lexique, ils leur remettent les clés qui leur donnent accès au moteur même de la langue, ce qui, nous en sommes convaincu, leur donnera envie de mieux en saisir les rouages.

BIBLIOGRAPHIE

- Andre, T. (1986). *Problem solving and education*. San Diego, CA: Academic Press.
- Antidote Prisme*, version 5 [Logiciel], Montréal, Druide informatique, 2004.
- Barrera-Vidal, A. (1992). « Peut-on, doit-on, valoriser les fautes de nos apprenants ». *Enjeux*, no 27, 54-60.
- Berthoud, A.-C. (1982) « Sur la relative fiabilité du discours métalinguistiques des apprenants ». *Encrages*, 8-9, 139-142.
- Bulman, F. (2003). *Le prépositionnaire : dictionnaire des verbes et des adjectifs pouvant être suivis d'une préposition*. Québec : L'instant même.
- Bureau, C. (1985). *Le français écrit au secondaire : une enquête et ses implications pédagogiques*. Québec : Conseil de la langue française.
- Cajolet-Laganière, H. et Martel, P. (1995). *La qualité de la langue au Québec*. Québec : Institut québécois de recherche sur la culture, Coll. Diagnostic.
- Catach, N. (1995). *L'Orthographe française*, 3^e édition. Paris : Nathan Université.
- Chaudenson, R., Mougeon, R. et Beniak, É. (1993). *Vers une approche panlectale de la variation du français*. Paris : Diffusion Didier-Érudition.
- Cicurel, F. (1985). *Parole sur parole ou Le métalangage dans la classe de langue*. Paris : Clé international, Coll. Didactiques des langues étrangères.
- Cogis, D. et Ros, M. (2003). « Les verbalisations métagraphiques : un outil didactique en orthographe ? », *Dossiers des sciences de l'éducation de l'université de Toulouse-le Mirail*, 9, 89-98.
- Conseil de la langue française. (1998). *Maîtriser la langue pour assurer son avenir*. Québec : Conseil de la langue française.
- Corbeil, J.-C. (1980). « Les choix linguistiques », dans *Actes du colloque « La qualité de la langue... après la loi 101 »*. Québec : Conseil de la langue française.
- Dallaire, É. (1997). *Évolution du discours institutionnel tenu sur la langue, la norme et la grammaire depuis le rapport Parent*. "Mémoire de maîtrise du 2^{ème} cycle non publiée" Université du Québec à Montréal, Montréal.
- David, J. et Jaffré, J.-P. (1997). « Le rôle de l'autre dans les procédures métagraphiques. » *Recherches, revue de didactique et de pédagogie du français*, 26, 155-168.
- Denance, M. (1999). « L'évaluation des fautes de vocabulaire à l'évaluation de l'épreuve uniforme ». *Correspondance*, vol.4, no4.
<http://www.ccdmd.qc.ca/correspo/Corr4-4/Denance.html>

De Villers, M.-É. (1999). « Usages lexicaux propres au français du Québec ». *Correspondance*, vol.4, no4.
<http://www.ccdmd.qc.ca/correspo/Corr4-4/Villers.html>

De Villers, M.-É. (2003). *Le Multidictionnaire de la langue française*. Montréal : Québec Amérique.

Dolz, J. (1998) « Activités métanlangagières et enseignement du français ». Dans Dolz, J. et Meyer, J.-C. (dir). *Activités métalangagières et enseignement du français : actes des journées d'étude en didactique du français (Cartigny, 28 février-1 mars 1997)*, Bern ; Berlin ; Frankfurt/M. ; New-York ; Paris ; Wien : Lang.

Dreyfus, M. (2004, août). *Pratiques et représentations de l'enseignement du lexique à l'école primaire*. Communication présentée au 9^e colloque de l'Association Internationale pour la Recherche en Didactique du Français, Québec.

Dubois, J. (1994). *Dictionnaire de linguistique et des sciences du langage*. Paris : Larousse, Coll. Trésors du français.

Ducrot, O. et Schaeffer, J.-M. (1995). *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris : Éditions du Seuil, 311-320.

Forest, C. et Boudreau, D. (1999). *Dictionnaire des anglicismes Le Colpron*. Laval, Québec : Groupe Beauchemin, éditeur ltée.

Fortier, G. (1994) « Propos sur le vocabulaire ». *Québec français*, no 92, 24-27.

Gagné, E.D. (1985). *The cognitive psychology of school learning*. Boston, MA: Little, Brown and company.

Glover, J. A., Ronning, R. R. & Bruning, R. H. (1990). *Cognitive psychology for teachers*. New-York : MacMillan.

Gombert, J. É. (1990). *Le développement métalinguistique*. Paris : Presse universitaires de France, Coll. Psychologie d'aujourd'hui.

Grevisse, Maurice. (2004). *Quelle préposition ? : cinquième édition*. Paris : Duculot, Coll. Entre Guillemets.

Groupe DIEPE. (s.d.). *Savoir écrire au secondaire. Étude comparative auprès de quatre populations francophones d'Europe et d'Amérique*, s.l., De Boeck Université.

Honveau, R. (1995) « Variation linguistique et norme : variables et erreurs graphiques ». *Liaison-HESO, De la variation : écarts, erreurs, normes*. no double 25-26, 53-64.

Jaffré, J.-P. (1995). Compétence orthographique et acquisition. Dans D. Ducard, R. Honvault et Jaffré, J.-P. *L'Orthographe en trois dimensions*. Paris : Nathan pédagogie, 93-158.

Jaffré, J.-P. et Ducard, D. (1996). « Approches génétiques et productions graphiques ». *Études de linguistique appliquée*, 101, 87-98.

Jaffré, J.-P. (1997). « Gestion et acquisition de l'orthographe ». *Revue française de linguistique appliquée*, vol. II-2, 61-70.

Jaffré, J.-P. et David, J. (1999). « Le nombre : essai d'analyse génétique ». *Langue française*, 124, 7-22.

Jakobson, R. (1963). *Essai de linguistique générale*. Paris : Éditions de Minuit.

Lafontaine, L. et Legros, C. (1995). « Profils linguistiques, cognitifs et motivationnels d'étudiants du postsecondaire faibles en français écrit ». *Revue des Sciences de l'Éducation*, vol.21, no1, 121-144.

Leeman-Bouix, D. (1994). *Les fautes de français existent-elles ?* Paris : du Seuil.

Lefrançois, P. (2003). « Comment résout-on les problèmes orthographiques liés aux homophones au primaire? ». *Dossiers des sciences de l'éducation de l'université de Toulouse-le Mirail*, 9, 67-76.

Lefrançois, P. et Montesinos-Gelet, I. (à paraître). « Les commentaires métagraphiques, reflet et soutien du développement linguistique ».

Lefrançois, P. (2004). « Les stratégies de résolution de problèmes orthographiques d'étudiants à l'université ». Dans C. Barré-De Miniac, C. Brissaud et M. Rispail (dir.), *La littératie : conceptions théoriques et pratiques de l'enseignement de la lecture-écriture*. Paris : l'Harmattan, 233-244.

Lefrançois, P.; Laurier, M. D.; Lazure, R. et Claing, R. (2005). *Évaluation de l'efficacité des mesures visant l'amélioration du français écrit du primaire à l'université*. Rapport de recherche. Montréal : Université de Montréal.

Legros, C. et Roy, G.-R. (1995). « Du discours métalinguistique tenu par les étudiants du postsecondaire relativement aux accords de genre et de nombre ». *Cahier de la recherche en éducation*, 2(2), 227-257.

Le Petit Robert. Dictionnaire de la langue française, sous la direction de Josette Rey-Debove et d'Alain Rey. (2002). Paris : Le Robert.

Le Petit Robert. Dictionnaire alphabétique et analogique de la langue française, version électronique 2.1, sous la direction de Josette Rey-Debove et d'Alain Rey (2001). Paris : Le Robert.

Lépine, F. (1995). « Bilan des tests de français à l'admission aux universités québécoises (1987-1994) ». *Revue des sciences de l'éducation*, vol.21, no1, 17-36.

Martel, P. et Cajolet-Laganière, H. (1996). *Le français québécois ; usages, standard et aménagement*. Québec : Institut québécois de recherche sur la culture, Coll. Diagnostic.

- Mel'čuk, I., Zholkovsky, A. (1984). *Explanatory Combinatorial Dictionary of Modern Russian*. Vienne: Wiener Slawistischer Almanach.
- Mel'čuk, I., (avec Arbatchewsky-Jumarie, N., Elnitsky, L., Iordanskaja, L. et Lessard, A.; réd. Clas, A.). (1984). *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques I*. Montréal : Presses de l'Université de Montréal.
- Mel'čuk, I., (avec Arbatchewsky-Jumarie, N., Dagenais, L., Elnitsky, L., Iordanskaja, L., Lefebvre, M.-N., et Mantha, S.; réd. Clas, A.). (1988). *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques II*. Montréal : Presses de l'Université de Montréal.
- Mel'čuk, I., (avec Arbatchewsky-Jumarie, N., Iordanskaja, L., et Mantha, S.; réd. Clas, A.). (1992). *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques III*. Montréal : Presses de l'Université de Montréal.
- Mel'čuk, I., (avec Arbatchewsky-Jumarie, N., Iordanskaja, L., Mantha, S. et Polguère, A.). (1999). *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques IV*. Montréal : Presses de l'Université de Montréal.
- Mel'čuk, I., Clas, A., et Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve, France : Duculot.
- Mel'čuk, I. (1997). *Vers une linguistique Sens-Texte : Leçon inaugurale faite le Vendredi 10 janvier 1997*, Paris, Collège de France : Chaire internationale.
- Mel'čuk, I. (2003). « Collocations dans le dictionnaire ». Dans Szende, T. (dir.), *Les écarts culturels dans les dictionnaires bilingues*. Paris : Champion, 19-64.
- Ministère de l'Éducation. (1997). *Résultats aux épreuves uniques de 1996 par commission scolaire et par établissement privé et diplomation par commission scolaire*, Québec.
- Ministère de l'Éducation. Direction générale des affaires universitaires et collégiales, Direction de l'enseignement collégial. (2004). *Résultats aux épreuves de langue d'enseignement et littérature. Année scolaire 2003-2004*. Consulté le 19 août 2005 sur le site Web du ministère de l'Éducation, du Loisir et du Sport.
http://www.mels.gouv.qc.ca/ens-sup/ens-coll/Eprv_uniforme/statistiques.asp
- Moffet, J.-D. (1999). « La correction et l'épreuve uniforme de français : un « fardeau » à partager sur une même longueur d'onde ». *Correspondance*, vol. 4, no 3.
<http://www.ccdmd.qc.ca/correspo/Corr4-3/Moffet.html>
- Montésinos-Gelet, I. et Morin, M.-F. (2001). « S'approcher de la norme orthographique en 1^{re} année du primaire : qu'en est-il de la pluralité des conceptions linguistiques ? ». *Archives de Psychologie*, 69, 159-176.
- Morin, M.-F. (2002). *Le développement des habiletés orthographiques chez des sujets francophones entre la fin de la maternelle et de la première année du primaire*. "Thèse de doctorat non publiée". Université Laval, Québec.

Patris, J. et Vansnick, N. (1992). « Comment aborder les fautes lexicales en classe ». *Enjeux*, no 26, 46-55.

Petiot, G. (2000). *Grammaire et linguistique*. Paris : Armand-Colin, Coll. Campus linguistique.

Picoche, J. (1992). *Précis de lexicologie française ; l'étude et l'enseignement du vocabulaire*. Paris : Nathan, Coll. Fac. Linguistique.

Polguère, A. (2003). *Lexicologie et sémantique lexicale : Notions fondamentales*. Montréal : Les Presses de l'Université de Montréal, Coll. Paramètres.

Polguère, A. (à paraître). « "Soleil insoutenable" et "chaleur de plomb" : le statut linguistique des greffes collocationnelles ». Dans L'Homme, M.-C. et Vandaele, S. (éd.), *Les approches lexicographiques et terminologiques sont-elles compatibles ?* [titre provisoire]. Ottawa : Presses de l'Université d'Ottawa.

Réseau pour la nouvelle orthographe du français (2004). *Vadémécum de l'orthographe recommandée (Le millepatte sur un nénufar)*. Belgique/France/Québec/Suisse : Publication du RENOUVO.

Rey-Debove, J. (1997). *Le métalangage*. Paris : Armand Colin/Masson.

Ros Dupont, M. (1997). « Les explications métagraphiques, une fenêtre ouverte sur l'acquisition de l'écrit. L'exemple de l'expression graphique du nombre ». *Liaisons HESO*, 29-30, 11-60.

Roy, G.-R., Boudreau, G., Viau, R. et Lafontaine, L. (1993). *Analyse de cas problèmes en maîtrise du français écrit aux ordres d'enseignement collégial et universitaire*. Dans L.-G. Bordeleau, M. Brabant, B. Cazabon, F. Desjardins et R. Leblanc (éd.), *Libérer la recherche en éducation – Actes du 3^e congrès des sciences de l'éducation de langue française au Canada*. Ottawa : Centre franco-ontarien de ressources pédagogiques, 85-101.

Roy, G.-R., Lafontaine, L. et Legros, C. (1995). *Le Savoir grammatical après treize ans de formation*. Sherbrooke : Éditions du CRP.

Roy, G.-R. (1995) « Résolution de problèmes d'ordre syntaxique par des étudiants du postsecondaire ». *Revue des Sciences de l'Éducation*, vol.21, no1, 167-195.

Saussure, F. de (1972). *Cours de linguistique générale*. Paris : Payot.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge : Cambridge University Press. Coll. Cambridge language education.

Simard, C. (1994). « Pour un enseignement plus systématique du lexique ». *Québec français*, no 92, 28-33.

Stubbs, M. (2001). *Words and phrases : Corpus studies of lexical semantics*. Oxford/Malden, Mass. : Blackwell, 3-23.

Table Éducation Montréal. (2002). *100 façons d'améliorer le français écrit des élèves*. Montréal : Table de concertation interordres montréalaise.

Tardif, J. (1992). *Pour un enseignement stratégique : L'apport de la psychologie cognitive*. Montréal : Logiques.

Tesnière, L. (1959). *Éléments de syntaxe structurale*. Paris : Klincksieck.

Van der Maren, J.M. (1996) *Méthodes de recherche pour l'éducation, 2e édition*. Montréal : PUM ; Bruxelles : De Boeck-Université.

Vermersch, Pierre. (1994). *L'entretien d'explicitation*. Paris : ESF Éditeur, Coll. Pédagogies.

ANNEXE I :
formulaire de
consentement

Formulaire de consentement

Je, soussigné(e), accepte de participer à la recherche intitulée *Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes lexicaux et analyse des stratégies de résolution de problèmes lexicaux*, menée par Dominic Ancil dans le cadre de sa maîtrise en didactique. Ce faisant, j'accepte que ma rédaction, faite pour le cours DID1010A, soit utilisée comme élément d'un corpus de textes servant à recenser des erreurs lexicales.

Je reconnais avoir reçu des informations satisfaisantes sur les objectifs de la recherche et l'usage des données qui seront recueillies.

Je sais que les données recueillies resteront confidentielles et ne seront utilisées que par le chercheur et que dans le cadre de la présente recherche. Je sais aussi que mon nom ne sera jamais mentionné dans le mémoire.

Nom : _____ Groupe : DID1010A-_____

Signature : _____ Date : _____

Informations relatives au profil linguistique :

1. Inscrivez votre/vos **langue(s) maternelle(s)** (celles que vous avez apprises à la maison).

2. Avez-vous **toujours** été scolarisé en français (c'est-à-dire depuis la première année du primaire jusqu'à aujourd'hui)?

OUI _____

NON _____ Détails : _____

Formulaire de consentement

Je, soussigné(e), accepte de participer à la recherche intitulée *Maîtrise du lexique chez les étudiants universitaires : typologie des problèmes lexicaux et analyse des stratégies de résolution de problèmes lexicaux*, menée par Dominic Ancil dans le cadre de sa maîtrise en didactique. Ce faisant, j'accepte d'être filmé(e) et enregistré(e) lors d'une tâche de révision de textes menée en équipe.

Je reconnais avoir reçu des informations satisfaisantes sur les objectifs de la recherche et l'usage des données qui seront recueillies.

Je sais que les données recueillies resteront confidentielles et ne seront utilisées que par le chercheur et que dans le cadre de la présente recherche. Je sais aussi que mon nom ne sera jamais mentionné dans le mémoire.

Nom : _____

Signature : _____ Date : _____

Date de naissance : ____ / ____ /19 ____

Études :

Programme d'études : _____

Université : _____

Année : _____

Informations relatives au profil linguistique :

1. Inscrire votre/vos **langue(s) maternelle(s)** (celles que vous avez apprises à la maison).

2. Avez-vous **toujours** été scolarisé en français (c'est-à-dire depuis la première année du primaire jusqu'à aujourd'hui)?

OUI _____

NON _____ Détails : _____

ANNEXE II :
test du deuxième
objectif

Première partie
La discipline dans les écoles

Peut-on affirmer que la discipline existe toujours dans les écoles du Québec ? Certains disent que plus le temps a passé, plus les générations sont devenues davantage permissives. Je crois pour ma part que, bien que le fouet n'est plus utilisé dans les écoles du Québec, d'autres formes de discipline sont aujourd'hui apparues.

De nos jours, la paresse attaque beaucoup plus les jeunes qu'autrefois, en particulier ceux du secondaire. En effet, ceux-ci sont plutôt portés de manquer leurs cours que d'étudier leurs leçons. Pour cette raison, les enseignants doivent employer leur autorité pour faire comprendre aux élèves l'importance des études. Faire en sorte que ces deux groupes corroborent est un réel défi, et les professeurs doivent parfois imposer des conséquences aux élèves. Mais plutôt que de mettre l'emphasis sur les punitions, l'enseignant devrait donner à l'élève l'opportunité d'apprendre par lui-même à appliquer un plus grand respect de soi et des autres. L'enseignant pourrait lui-même devenir un modèle de respect en s'intéressant davantage aux intérêts des jeunes, à leur réalité et aux différents idoles de l'heure. Ainsi, les élèves se sentiraient plus proches de leur enseignant et déprimeraient moins de devoir se rendre à l'école.

Pour en terminer avec la discipline, je désire mentionner qu'elle est essentielle au bon fonctionnement d'une école. Par contre, il est important que les enseignants comprennent que l'excès de discipline, comme un trop grand laxisme, peut affecter l'intérêt des jeunes envers les études. C'est donc au professeur de trouver le juste et bon milieu dans l'exercice de la discipline.

Deuxième partie
Phrases indépendantes

1. Le jeune homme **emmenait** toujours son calepin avec lui.

2. À l'époque, personne sauf moi n'entrait dans ma chambre. Un jour, j'ai décidé de **donner l'accès** à mon frère, qui était ravi.

3. Ce magnat de l'industrie textile avait trempé dans une histoire de **blanchissage d'argent**.

4. Les travaux ont dû être interrompus à **cause que** la main d'œuvre qualifiée préfère les chantiers du Sud, qui offrent de meilleures conditions.

5. Émile était très distrait ce matin-là; il **avait même monté** dans le mauvais autobus.

6. Il est étrange que des gens qui ont tous **transité** par le secondaire ne se souviennent que partiellement de cette période de leur cheminement scolaire.

7. Les journalistes ont publié des photos compromettantes de l'homme politique, **question de salir sa réputation.**

8. Le fait de donner l'occasion aux enfants de gérer leur temps dès la fin du primaire leur permet de **prendre la responsabilité de leurs études entre leurs mains.**

9. Un grand parc longeait **ma rue de résidence.**

10. L'inspecteur Massicotte tente de faire la lumière sur l'agression de la jeune femme. **Elle lui informe ses déplacements** pour le besoin de l'enquête.

11. Ces 92 pages nous **racontent** les péripéties d'une journée dans la vie d'un banquier londonien.

12. L'école ne doit pas être seule à s'occuper de l'éducation des jeunes. Je pense que les parents devraient **avoir beaucoup plus de discipline** à la maison avec leurs enfants.

13. J'adorais les après-midi passés chez mes grands-parents. Je me rendais au grenier et j'explorais ce qui s'y trouvait.

14. Tant qu'à moi, il vaudrait mieux imposer une taxe spéciale aux gros véhicules que pénaliser l'ensemble des automobilistes.

15. La routine a rapidement repris son cours et l'enthousiaste des enfants pour leur nouveau compagnon a beaucoup diminué.

**ANNEXE III :
consignes pour le
test du deuxième
objectif**

Consignes données aux sujets
pour le test du deuxième objectif

Le test que vous allez faire aujourd'hui comporte deux parties. Pour les deux sections, vous disposerez d'autant de temps que vous désirez.

Je vous demanderais de toujours lire à voix haute les textes qui vous sont soumis afin que je sache exactement où vous êtes rendus. Il est aussi important que vous verbalisiez l'ensemble de vos réflexions, peu importe si vous êtes convaincus ou non de vos explications. Expliquez toujours à votre collègue les changements que vous désirez apporter au texte en les justifiant autant que possible. Tous les passages modifiés devront être rayés sur le texte original et les modifications apportées sur la feuille au-dessus de l'extrait concerné. Tout ce que vous rayez doit l'être d'un trait simple afin que je puisse encore le lire.

Si je constate que vous ne verbalisez pas certains raisonnements ou que vous êtes bloqués à certains endroits, il se peut que j'intervienne en vous posant des questions. Si je vous interroge sur un élément spécifique du texte, il ne faut pas croire automatiquement qu'il s'agit d'une erreur; si vous ne le percevez pas comme un problème, il faut simplement me dire pourquoi.

Finalement, vous disposez d'un *Petit Robert* et d'un *Multidictionnaire* auxquels vous pouvez vous référer aussi souvent que vous le voulez.

Première partie : La discipline dans les écoles

Voici un texte d'une page qui porte sur la discipline dans les écoles. Ce texte a été rédigé à partir de rédactions faites par des étudiants universitaires. Il présente certains problèmes (je ne vous dis pas combien) et je vous demande de tenter de les repérer. Je ne vous dis pas non plus de quels types de problèmes il s'agit.

Vous devez donc lire le texte à voix haute et vous arrêter sur les items que vous trouvez problématiques. Il faut ensuite m'expliquer pourquoi ces éléments constituent des problèmes et quels changements vous pourriez apporter pour rendre correct l'extrait concerné.

Le présent test porte précisément sur la dimension linguistique du texte et non pas sur le contenu ou la valeur des arguments présentés. De plus, vous devez corriger le texte comme si vous étiez un enseignant; par conséquent, il faut modifier seulement ce que vous considérez comme des erreurs et non pas récrire complètement le texte comme vous l'auriez vous-mêmes rédigé.

Vous pouvez relire le texte plus d'une fois.

Deuxième partie : Phrases indépendantes

Voici maintenant quinze phrases dans lesquelles certains mots ou groupes de mots apparaissent en caractères gras; ce sont des éléments que je vous présente comme problématiques.

Vous devez lire chaque phrase à voix haute et spécifier si le passage en gras constitue véritablement un problème selon vous. S'il s'avère effectivement erroné, vous devez expliquer pourquoi et proposer une solution, comme pour la première partie du test.

Les passages qui ne sont pas en caractères gras ne présentent aucune erreur; je vous demande donc de ne pas les modifier.